

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:32:18 / Search time 31.6461 Seconds
(without alignments)
2832.275 Million cell updates/sec

Title: US-09-856-061-2
Perfect score: 2316
Sequence: 1 MTSQGNKRTKEGFDLRFQ.....QPLPLARLLTQYSSQALHE 435

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21.*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_minc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_protein:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriaph:*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	11 Q9JMJ3	Q9JMJ3 mus musculus
2	2305	99.5	435	11 Q9OZP2	Q9OZP2 mus musculus
3	1132	48.9	376	4 Q9P2U9	Q9P2U9 homo sapien
4	410.5	17.7	533	11 Q922M0	Q922M0 mus musculus
5	410	17.7	525	11 Q920L0	Q920L0 ratius norv
6	381	16.5	530	13 Q9DC07	Q9DC07 gallus gall
7	319	13.8	456	4 Q8WV28	Q8WV28 homo sapien
8	319	13.8	456	4 Q75498	Q75498 homo sapien
9	315.5	13.6	552	13 Q9YGC1	Q9YGC1 gallus gall
10	309.5	13.4	433	4 Q75499	Q75499 homo sapien
11	297.5	12.8	457	11 Q9QUN3	Q9QUN3 mus musculus
12	287.5	12.4	457	11 Q88504	Q88504 mus musculus
13	219.5	9.5	297	11 Q9DA13	Q9DA13 mus musculus
14	188.5	8.1	516	11 Q54737	Q54737 mus musculus
15	167.5	7.2	471	4 Q9UCX5	Q9UCX5 homo sapien
16	165.5	7.1	594	4 Q92529	Q92529 homo sapien

17	165.5	7.1	594	4 Q8TAP2	Q8TAP2 homo sapien
18	150.5	6.5	486	4 Q9UCX4	Q9UCX4 homo sapien
19	150.5	6.5	812	15 Q85466	Q85466 w73 sarcoma
20	143.5	6.2	662	5 Q9N3S5	Q9N3S5 caenorhabdi
21	141.5	6.1	474	11 Q61120	Q61120 mus musculus
22	141	6.1	728	4 Q9UPH9	Q9UPH9 homo sapien
23	139.5	6.0	286	5 Q95PX0	Q95PX0 caenorhabdi
24	139.5	6.0	398	5 Q95PW9	Q95PW9 caenorhabdi
25	138.5	6.0	335	4 Q96CZ8	Q96CZ8 homo sapien
26	138.5	6.0	926	4 Q9HOK1	Q9HOK1 homo sapien
27	137.5	5.9	594	11 Q70143	Q70143 ratius norv
28	136.5	5.9	477	11 Q70142	Q70142 ratius ralt
29	134	5.8	309	11 Q9OVZ0	Q9OVZ0 mus sp. shb
30	134	5.8	461	11 Q63789	Q63789 ratius norv
31	134	5.8	1599	5 Q18892	Q18892 caenorhabdi
32	133.5	5.8	217	11 Q9CX99	Q9CX99 mus musculus
33	132.5	5.7	904	5 Q9VMJ4	Q9VMJ4 drosophila
34	132.5	5.7	944	5 Q8T9B3	Q8T9B3 drosophila
35	130.5	5.6	722	13 Q8U0U2	Q8U0U2 xenopus lae
36	129	5.6	559	11 Q91252	Q91252 mus musculus
37	129	5.6	645	6 Q95JY5	Q95JY5 macaca fasc
38	128.5	5.5	112	11 P70305	P70305 mus musculus
39	128.5	5.5	393	2 Q93CA3	Q93CA3 treponema p
40	128.5	5.5	854	11 Q9QZS8	Q9QZS8 mus musculus
41	128.5	5.5	1430	11 Q8VHK2	Q8VHK2 ratius norv
42	128	5.5	623	6 Q95JL7	Q95JL7 macaca fasc
43	128	5.5	651	11 P97504	P97504 mus musculus
44	128	5.5	1160	11 Q9ESC8	Q9ESC8 mus musculus
45	128	5.5	4498	13 Q93291	Q93291 fugu rubrip

ALIGNMENTS

RESULT 1					
Q9JMJ3	PRELIMINARY:	PRT:	435 AA.		
AC Q9JMJ3;					
DT 01-OCT-2000 (TREMBLrel. 15, Created)					
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)					
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)					
DE MIST.					
GN CLINK OR MIST.					
OS Mus musculus (Mouse).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
OX NCBI_TaxID=10090;					
RN [1]					
RP SEQUENCE FROM N.A.					
RX MEDLINE-20208989; PubMed-10744659;					
RA Gotsuka R., Kanazashi H., Sasanna H., Fujimura Y., Hida Y.,					
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.,					
RT "A BASH/SLP-76-related adaptor protein MIST/CLINK involved in Ige					
RT receptor-mediated mast cell degranulation.";					
RL Int. Immunol. 12:573-580(2000).					
DR EMBL, AB021220; BAA96240.1; -					
DR HSSP: P23727; 2PNB.					
DR MGD: MGI:1351468; CLINK.					
DR InterPro: IPR000980; SH2.					
DR Pfam: PF00017; SH2; 1.					
DR PRINTS: PR00401; SH2DOMAIN.					
DR Prodom: PD000093; SH2; 1.					
DR SMART: SM00252; SH2; 1.					
DR PROSITE: PS0001; SH2; 1.					
SQ SEQUENCE 435 AA; 49513 MW; 4D1BD3E2F0C61ED6 CRC64;					
Query Match	100.0%;	Score 2316;	DB 11;	Length 435;	
Best Local Similarity	100.0%;	Pred. No. 8.3e-180;			
Matches 435;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;	
OY 1 MTSQGNKRTKEGFDLRFQNVSLKNRSPSLSSAKGRCRAVLEPLPHRRRLAGVPGG 60					
DB 1 MTSQGNKRTKEGFDLRFQNVSLKNRSPSLSSAKGRCRAVLEPLPHRRRLAGVPGG 60					

Qy	61	EKSNNSNDYEDBPBQOLKAWMSMILPARPJQSEYADTRYFQDMMAPLILPPKASVST	120
Db	61	EKCSNNNDYEDBPBQOLKAWMSMILPARPJQSEYADTRYFQDMMAPLILPPKASVST	120
Qy	121	ERQTRDVMTOLEEDKPTFDKVSQSPFKGKYYIKTLPRLPPRPATILTPKKYQPLPPA	180
Db	121	ERQTRDVMTOLEEDKPTFDKVSQSPFKGKYYIKTLPRLPPRPATILTPKKYQPLPPA	180
Qy	181	PEESSAYFAKPTFPYQVRQPRORSKADFSRYLGAEEESHNOOTKPRESSCSSNOMTQKS	240
Db	181	PEESSAYFAKPTFPYQVRQPRORSKADFSRYLGAEEESHNOOTKPRESSCSSNOMTQKS	240
Qy	241	PPAIASSSYMGKHSIQARDHTGSMQHCPAORCOAAASHSPRMILPYENTNSEKPDPTKPD	300
Db	241	PPAIASSSYMGKHSIQARDHTGSMQHCPAORCOAAASHSPRMILPYENTNSEKPDPTKPD	300
Qy	301	EKDWNQNMWTYIGEXSRQAVEDVLKKNKDGFLVIRBDSSTKSKEPYLVYFYCNKYNNK	360
Db	301	EKDWNQNMWTYIGEXSRQAVEDVLKKNKDGFLVIRBDSSTKSKEPYLVYFYCNKYNNK	360
Qy	361	IRFLESNQOFAIGTGLRGNEFMFDSVEDIIEHYTYFPIILLIDGDKAKARRKQCLTQPLPL	420
Db	361	IRFLESNQOFAIGTGLRGNEFMFDSVEDIIEHYTYFPIILLIDGDKAKARRKQCLTQPLPL	420
Qy	421	ARLLLTQYSQALHE	435
Db	421	ARLLLTQYSQALHE	435

RESULT 2
Q9QZE2
ID Q9QZE2 PRELIMINARY; PRT; 435 AA

DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE CLNK.
GN CLNK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euteleia; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20029811; PubMed=10562326;
RA Cao M.Y., Davidson D., Yu J., Latour S., Velliette A.;
RT "Clnk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
RT Stimulated Hemopoietic Cells.";
RL J. Exp. Med. 190:1527-1534(1999).
DR EMBL: AF187819; AF14299.1; -.
DR HSSP: P23727; 2PNB
DR MGD; MGI:1351468; CLNK.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PS0001; SH2; 1.
SQ SEQUENCE 435 AA; 49492 MW; 5CD27EC971FCD8A5 CRC64;

Query Match	99.58	Score 2305	DB 11	Length 435
Best Local Similarly	99.58	Pred. No. 6.5e-179		
Matches 433	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

Oy	1	MTSGGNKRTTEEGGCDLRFQVNSLLTKNRSPSTSSAAGRCRAVLPEPLPHDRRLNLAAPGC	60
Db	1	MTSGGNKRTTEGGCDLRFQVNSLLTKNRSPSTSSAAGRCRAVLPEPLPHDRRLNLAAPGC	60
Oy	61	EKCNSNDYEDPEFOLLKAWESMKILPARPIQSEYADTRYFDQMMEARPLLPPKASVST	120
Db	61	EKCNSNDYEDPEFOLLKAWESMKITLPARPIQSEYADTRYFDQMMEARPLLPPKASVST	120

Qy	121	ER0TRDVRMT0LEEDYDKPTFKDVSNS0RPFKGKYYTKIKTLP,PPRPAITLP,PKYQYLP,PA	180
Db	121	ER0TIDVYMTHTLEEEDKPTFKDVSNS0RPFKGKYYTKIKTLP,PPRPAITLP,PKYQYLP,PA	180
Qy	181	PPRESSAFAKPTFPPEYQORPGRORSKADPSRYLGAEEESHNDTKPPSSCPSSNOMTOKS	240
Db	181	PPRESSAFAKPTFPPEYQORPGRORSKADPSRYLGAEEESHNDTKPPSSCPSSNOMTOKS	240
Qy	241	PPAISSSSYMGKHSIQARDITGSMOHCAPORCOAAASHSPMLPYENTNSEKPPDPTKPD	300
Db	241	PPAISSSSYMGKHSIQARDITGSMOHCAPORCOAAASHSPMLPYENTNSEKPPDPTKPD	300
Qy	301	EKDWNQNMWYIGEXYSRQAVEDEVLMKENKDGTFILVRDOSTSKAEPYLVVYFNKNKYNNK	360
Db	301	EKDWNQNMWYIGEXYSRQAVEDEVLMKENKDGTFILVRDOSTSKAEPYLVVYFNKNKYNNK	360
Qy	361	IRFLESNOQFALGTGLRGNEFMFDSVEDIIEHYTYFPIILLIDGKDKAARRKQCYLT0PPL	420
Db	361	IRFLESNOQFALGTGLRGNEFMFDSVEDIIEHYTYFPIILLIDGKDKAARRKQCYLT0PPL	420
Qy	421	ARLLITQYSSQALHE	435
Db	421	ARLLITQYSSQALHE	435

RESULT 3
Q9P2U9
ID Q9P2U9 PRELIMINARY; PRT; 376 AA

```

DT      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      MIST (Fragment) .
GN      MIST.
OS      Homo sapiens (Human) .
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX      NCBI_TaxID=9606;
RX      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20208989; PubMed=10744659;
RA      GOLTSUKA R., KANAZASHI H., SASANUMA H., FUJIMURA Y., HIDAKA Y.,
RA      TATSUNO A., RA C., HAYASHI K., KITAMURA D.;
RT      "A BASH/SLP-76-related adaptor protein MIST/Cink involved in Ige
RT      receptor-mediated mast cell degranulation." ;
RL      Int. Immunol. 12:573-580(2000) .
DR      EMBL; AB032369; BAA96241.1; .
DR      HSSP; P23727; 2PNB.
DR      InterPro; IPR000980; SH2.
DR      Pfam; PF00017; SH2; 1.
DR      PRINTS; PR00401; SH2DOMAIN.
DR      ProDom; PD000093; SH2; 1.
DR      SMART; SM00252; SH2; 1.
DR      PROSITE; PSS0001; SH2; 1.
FT      NON_TER      1      1
SQ      SEQUENCE      376 AA; 43542 MW; 389421B629B028B2 CRC64;

```

Query Match	48.98	Score 1132	DB 4	Length 376
Best Local Similarity	60.98	Pred. No. 8.2e-84		
Matches 229	Conservative 42	Mismatches 101	Indels 4	Gaps 3

```

QY 19 FQVSLKLNRSMPSTSSAKGCGRAVLDPDRHRLNLAGVGGKGNKNNDYEDPEPOLLK 78
Db 1 FQNSFLPNKRSWPIINSTGYQRMKNPLDMERNFAAVLDGAKGSHDDYDDPELIMEE 60
QY 79 AWPSSKILPARIPIESEAADRYEQDMEAFLPLPKASVSTFRQTDVAMTOLLEEVDR 138
Db 61 TWQSKILPARIPIESEAADHYFKVAMDPLPLDPTSTISIQPTMNG-TLELVNDR 119
QY 139 TFKDQSRQRFKGFYTKINKTPLRPPRAITLPLKKYQOLPDPAPBESSAYFADKPTFPEV 198

```

```
Db 120 ISRDVRSQNIKGDASVKNKIPPLPPRLITPKKYQPLP-EPSSRPLSQRTPEV 178
Qy 199 QGPRORSKADFSRYLGAEEESHQTKPESSCPSSNONTKSPPAIASSYMPGHSIOA 258
Db 179 QGMPQISILRDLSEVLEAEKYPHNOQRKESTHLENQTOEIPLAISSSSFTSNHVSQN 238
Qy 259 RDHTSMQHCPAQRQQAASHSP--RMLPYENTSEKPDPTKPDKDVQWONWYIGESR 316
Db 239 RDHRGMOPCSPQRCQPAQSCSPHENILPYKYSWPPPKRDKRDVQHNWYIGESR 298
Qy 317 QAVEDVLKMKNDGTFIVRDCSTKSKAEYVLVFGYGNKYVWKIRFLESNOQFALGTGL 376
Db 299 QAVEAPFKENKDGSLFVRDCSTKSKKEEYVLAVYENKYNVWKIRFLERNOQFALGTGL 358
Qy 377 RGENEMFSDVEDITIEHY 392
Db 359 RGDEKFDSDVEDITIEHY 374

RESULT 4
Q922M0 PRELIMINARY: PRT: 533 AA.
AC Q922M0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lymphocyte cytosolic protein 2.
GN LCP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC006948; AA06948.1;
DR MGD; MGI:1321402; LCP2.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR PROSITE; P55001; SH2; 1.
SQ SEQUENCE 533 AA; 60238 MW; 50AEA025EF0DAD01 CRC64;

Query Match 17.7%; Score 410.5; DB 11; Length 533;
Best Local Similarity 29.7%; Pred. No. 4,4e-25;
Matches 148; Conservative 50; Mismatches 176; Indels 125; Gaps 20;

Qy 14 FGDARFQNSL-----KNRSMPLSSAGRCRAVLEPLPDHRRNLG----- 56
Db 57 FPLRLPILSKLSQDINKNERSITRKQIRPLEETESHEDDGWGFEDDYESP 116
Qy 57 --VPGGKCSNNDYEDP--EFOLL-----KAMPS-----MKILPARPIQ--ESEX 96
Db 117 DDDPDE---DDGDYSPNEEQALVDADADYPPPSNNEALQSLIPNSFHTNSMY 173
Qy 97 ADTRRYQDMMEAPLLP--PKASY----- 118
Db 174 IDRPPLGKYSQPPVPLRKPALPPLPTGRNHSPLSPPHNHEEPSRSNNKTAKIPAP 233
Qy 119 STERQTRDVMTQLEEDV-----KPTFDYRSQRFKGYTKIKTKPLPPPAITL 170
Db 234 SIDRSTRPLDRSLAPLDRPFLGKKRPSDAPLGRHPLKTIKPLP--PAMDR 290
Qy 171 PKTYQPLPAPPESSAYFAFKPTFFEVQGRORSKADSRVLGAEEESHQTKPESSC 230
Db 291 HERHERLGPVTR-----KPPVPRHGRGPPRRE-----NDDDVHQRPLPQSL 334
Qy 231 PSSNONTKSPPAIASS-----SYMPGHSIOARDHTGSMQHC--PAQCOQAASHSP- 281
Db 335 PSMSSNTFPKRSVOPSSKNTFPLAHMPGARS--ESNIGQOQSALPPTYSQGGANRPL 391
Qy 282 ----RMLPYENTSEKPDPTKPDKDV--WQNEWYIGESRQAVEDVLKMKNDGTFIVRD 336
```

```
Db 392 RSEGNLPLPVPN--RPPSPGEEETPLDEEMVSYITTRAEALAKINKDQGLFVRD 449
Qy 337 CSTSKAEYVLVFGYGNKYVWKIRFLESNOQFALGTGLRGENMFSDVEDITIEHYTP 396
Db 450 SSKTKANNPYVLVLYKDKYVNIQIRYQESQVYLLGTGLGKEDFLSVSDIYFRKMP 509
Qy 397 ILLIDGKDKAARRKQCYLT 415
Db 510 ILLIDGKNRGS--RYOCTLT 527

RESULT 5
Q920L0 PRELIMINARY: PRT: 525 AA.
AC Q920L0;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE SLP-76 adaptor protein.
GN SLP76.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Fujii Y., Goitshuka G., Wakahara S., Hara T., Nakao T.;
RT "Molecular cloning of SLP-76, a 76-kDa adaptor protein in RBL-2H3."
RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AB072980; BAB71779.1;
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR PROSITE; P55001; SH2; 1.
SQ SEQUENCE 525 AA; 59521 MW; 0B1CDD40DABE7372 CRC64;

Query Match 17.7%; Score 410; DB 11; Length 525;
Best Local Similarity 31.9%; Pred. No. 4,8e-25;
Matches 127; Conservative 51; Mismatches 148; Indels 72; Gaps 16;

Qy 68 DYEDPEFOLKAMPKMLPAPPI--QSEYADTRYQDMMEAP-----LLLP--- 114
Db 144 DYPPPSNDEEALQN--SILPAKFPPTNSMYIDRPTGKVSQPPVPPQRPMAALPPLPT 202
Qy 115 -----KASYTERQTRDVMTQLEEDVDFQVDSQRRKGKTKINKTPL---- 161
Db 203 GRNHSNHEEPSRNHNKTKAKLPAPSIDRSTKPLD--NSLA-----PPLDREPLGKK 254
Qy 162 PPPRPAITLPPKKYQ---LPPAP-----EESAYFAPKPTFFEVQGRORSKADF 210
Db 255 PSDKPSAPLGRHPLKTIKPLRPAMDRIHRNRIRIPLTARKKPPVRHGRGPPRDR---- 310
Qy 211 SRVLGAEEESHQTKPESSCPSSNON-----TKSPPAIASSYMPGHSIOARDHTGS 264
Db 311 ----NDEDVHQRPLPSPSNTFPRSRSKPSKNTFPLPHMGALS--ESNIG 363
Qy 265 MQHC--PAQCOQAASHSP-----RMLPYENTSEKPDPTKPEKDVQWONWYIGESRQ 317
Db 364 QQSNASLPSYTSQGPSRPVRNEVKNLPLVPNRPOP--PSPGEEESPLDEEMVSYITTR 422
Qy 318 AVEDVLKMKNDGTFIVRDCSTKSKAEYVLVFGYGNKYVWKIRFLESNOQFALGTGLR 377
Db 423 EAELAKINKDQGLFVRDCSTKSKTKVNNPYVLVLYKDKYVNIQIRYQESQVYLLGTGLR 482
Qy 378 GNEMFSDVEDITIEHYTPILLIDGDKAARRKQCYLT 415
Db 483 GKEDFLSVSDIYFRKMPILLIDGKNRGS--RYOCTLT 519

RESULT 6
Q9D6G7 PRELIMINARY: PRT: 530 AA.
ID Q9D6G7
```

```

AC 09D607;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE SLP-76 adaptor protein.
CN SLP76.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_Taxid=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=THYMUS;
RX MEDLINE=20451093; PubMed=10993915;
RA Ishii M., Kurosaki M., Inabe K., Chan A.C., Sugamura K., Kurosaki T.;
RT "Involvement of LAT, Gads, and Grb2 in Compartmentation of SLP-76 to
RT the Plasma Membrane."
RL J. Exp. Med. 192:847-856(2000).
DR EMBL; AF226988; AAC18493.1; -
DR HSSP; P12931; 1SHD.
DR Interpro; IPR001660; SAM.
DR Interpro; IPR000980; SH2.
DR Pfam; PF00017; SH2.1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2.1.
DR SMART; SM00454; SAM; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
SQ SEQUENCE 530 AA; 59580 MW; CF6CBCC8D79909A CRC64;

Query Match 16.5%; Score 381; DB 13; Length 530;
Best Local Similarity 29.8%; Pred. No. 1,1e-22;
Matches 123; Conservative 54; Mismatches 151; Indels 85; Gaps 16;

QY 61 EKCSNND-YEDPEFOLKAMPKMLPAPPI-QESYATRTFQDMMEAPLLPPKASV 118
DB 136 EEAHSDSGTEPPPSNNDSEAHNV-IFPAKSLANNIDYIDRPTSRSSHQPPVPPORGP 194
QY 119 STERQTRDVMTQLEEVDRKPTFKDVSORFKGFKYTKINKTPLPAPPAITL----- 170
DB 195 SPAPASFGGASL-----PAFPLPENNDNRNK-----PKPPAPSIDRSTKPPDLR 242
QY 171 -----PK-----YOLPPAPPESSAYFAKPFPPE- 197
DB 243 LGPPERENVPGRKPGHPEKLLTPOLRALGEOLAMMPKPPVPPSYERGENSPLEKQI 302
QY 198 -VGRPGRSAKDFSRVLGAEEESH--HOTKPESSCPSSNQNT-----OKSPALIASSY 249
DB 303 PVAGMAQOKRPE-----EEDHITQRAVFOISLPYSSNTFPKSKIKAPKPGSNS- 354
QY 250 MEGKHSIQANDHTGSMOHCPSAQCQAASHSPM-----LPYENTNSEKPDPTKPE 301
DB 355 IPGAESARSLASAGSL---PPRPPLGNNSRSPRGTAADLRPLRPIPSRQAHQNTNEEDE 411
QY 302 KDVWONMWYGEYSRQAVEDVLMKENKDGTFILVDCSTSKAPYLVVYFGKVVNWKI 361
DB 412 -DLENDEMYAVYSRPAEALAKKINKDGFELVDSSTKTTTPYLVMLYKXKVNIQI 470
QY 362 RFLESNOQFALGTGLNGENMFDSVEDIIEHYTFPIILLIDGKKAARKQCYL 414
DB 471 RYQEOQDTYLLGTGLGKEDFSVAHIIDVFOFTPLILLIDGKDRGS-RNQCVL 522

RESULT 7
08WV28 PRELIMINARY; PRT; 456 AA.
AC 08WV28;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE B-cell linker.
RX Homo sapiens (Human).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=TONSIL;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC018906; AAH18906.1; -
DR Interpro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
SQ SEQUENCE 456 AA; 50465 MW; ED6D42A035D1792 CRC64;

Query Match 13.8%; Score 319; DB 4; Length 456;
Best Local Similarity 25.0%; Pred. No. 9,7e-18;
Matches 99; Conservative 58; Mismatches 139; Indels 100; Gaps 12;

QY 64 NSNNYEDPEFOLKAMPKMLPAPPIQESYADTRFYQDMMEAP---LLPPKASVST 120
DB 91 NADDSYEPPVE---QETRPVHPALPFARGEYIDNRSSQ--RHSPFSKTLPSKPSWPS 144
QY 121 ERQTRDVMTQLEEVDRKPTFKDVSORFKGFKYTKINKTPLPAPPAITLPPKYYQPLPA 180
DB 145 EKARLTSTLALALOKPY-----PPKRGLELDADYVVP 182
QY 181 PRESSAYFAKPTFPVGRPGRSAKDFSRVLGAEEESHQTKPESSCPSSNQNT--- 217
DB 183 EDNDENYIHPTESSSPPEKAPM-----VNRSTKNSSTPPAPSPETAGS 226
QY 238 -----OKSPALIASSYMP--GKH-----STQARDHTGSM---QHCPSAQCQA- 276
DB 227 RNSGAMETSSPP-APSPPLPRAKGPPTPLKTPVAAQQMASSVCEKPTPAERHRGSS 285
QY 277 ---ASHSPMLPYENTNSEKPPD-----TKPDKDVQNE 308
DB 286 HREQAVQSPVPPPAQNIHQKPIPLRFTEGGNPTVDGFLPSSTSSSTISQDAGVLCRP 345
QY 309 WYIGEYSRQAVEDVLMKENKDGTFILVDCSTSKAPYLVVYFGKVVNWKIRFLESNQ 368
DB 346 WTAGACDRKSAEALHRSNDSGFLIRKSSGHDSPQYLVVFNKRVYNIIVRFEATK 405
QY 369 QFALGTGLNGENMFDSVEDIIEHYTFPIILLIDGKD 404
DB 406 QVALGKRKNGEEFYSAELIRNHQSPVLVLDSON 441

RESULT 8
075498 PRELIMINARY; PRT; 456 AA.
AC 075498;
DT 01-NOV-1998 (Tremblrel. 08, Created)
DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)
DT 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE B cell linker protein BLNK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
RT receptor activation."
RL J. Biol. Chem. 272:27362-27368(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96361304; PubMed=9697839;
RA Fu C., Turck C.W., Kurosaki T., Chan A.C.;

```


RT "BLNK: a central linker protein in B cell activation."
RL Immunity 9:93-103(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE-20050956; PubMed-10583958;
RA Mineishi Y., Rohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
RA Campana D., Chan A.C., Conley M.E.;
RT "An essential role for BLNK in human B cell development."
RL Science 286:1954-1957(1999).
DR EMBL; AF068180; AAC39936.1; -
DR EMBL; AF180756; AAF20382.1; -
DR EMBL; AF180740; AAF20382.1; JOINED.
DR EMBL; AF180741; AAF20382.1; JOINED.
DR EMBL; AF180743; AAF20382.1; JOINED.
DR EMBL; AF180743; AAF20382.1; JOINED.
DR EMBL; AF180744; AAF20382.1; JOINED.
DR EMBL; AF180745; AAF20382.1; JOINED.
DR EMBL; AF180746; AAF20382.1; JOINED.
DR EMBL; AF180747; AAF20382.1; JOINED.
DR EMBL; AF180748; AAF20382.1; JOINED.
DR EMBL; AF180750; AAF20382.1; JOINED.
DR EMBL; AF180751; AAF20382.1; JOINED.
DR EMBL; AF180752; AAF20382.1; JOINED.
DR EMBL; AF180753; AAF20382.1; JOINED.
DR EMBL; AF180754; AAF20382.1; JOINED.
DR EMBL; AF180755; AAF20382.1; JOINED.
DR HSSP; P29354; 1BMB.
DR Interpro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
SQ SEQUENCE 456 AA; 50466 MW; 95FD5485D03D397 CRC64;

Query Match 13.8%; Score 319; DB 4; Length 456;
Best Local Similarity 25.0%; Pred. No. 9,7e-18;
Matches 99; Conservative 58; Mismatches 139; Indels 100; Gaps 12;

QY 64 NSNNDDYDPEFOLKAMPMSKILPARIQSEYADTRYFQDMMEAP---LILPKKASYST 120
DB 91 NADSYEPPEVE---OETRPVHPALPFAREXYIDNRSSQ---RHSPPFSKTLPSKPSWPS 144
QY 121 ERQTRDVRMTQLEVDKPTFKDVSQRFKGYKTKIKTKYLPRLPRPAITLPLKXQYLPRA 180
DB 145 EKARLTSTLPALALQRPV-----PRKPGLEDEADYVVPV 182
QY 181 PPESSAYFAKPPFPEVQGRPRORSADFSRVLGAEESHQTKPPESSCPSSNQNT--- 237
DB 183 EDNENNTIHPTESSPPPEKAPM-----VNRSTKPNSTTPASPGTASG 226
QY 238 -----OKSPPAIASSYMP--GKH-----SIQARDHTGSM--QHCPRACQQA-- 276
DB 227 RNSGAWETKSPPP--AASPLPRAKKPTPLKTPPVASQNAASSVCEKPIPARHRGSS 285
QY 277 ---ASHSPRLPYENNNSEKDP-----TKDEKDVQWNE 308
DB 286 HROFAVOSVFPRAOKOIHQRPILPRLPTEGNGPTVDGRLPFSFSSNSTISQDEAGVLCKP 345
QY 309 WYIGEYSROAVEDYLMKNDGTFLVRDCSTKSKAEPPVLYVFGNKYVNVKIRFLESNQ 368
DB 346 WYACACRKSMEALHNSKNKGSFLIRKSSGHDSCQRYTLVVFTRKRYNIPVFIETATK 405
QY 369 QFALGTGLRGNEMFDSVEDIIEHTYTPILLIDGKD 404
DB 406 QYALGRKKGEEYFGVAEIIRNHQHSPLVLIDSON 441

RESULT 9
QYGC1 PRELIMINARY; PRT; 552 AA.
AC QYGC1;
DT 01-MAY-1999 (TRIMBLrel. 10, Created)

DT 01-MAY-1999 (TRIMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TRIMBLrel. 20, Last annotation update)
DE B cell linker protein BLNK.
GN BLNK OR BASH.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinoptera; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99146381; PubMed-10023776;
RA Ishiai M., Kurosaki M., Pappu R., Okawa K., Ronko I., Fu C.,
RA Shibata M., Iwamatsu A., Chan A.C., Kurosaki T.;
RT "BLNK required for coupling Syk to PLC gamma 2 and Rac1-JNK in B
RT cells."
RL Immunity 10:117-125(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99049801; PubMed-9834055;
RA Gotsuka R., Fujimura Y., Mamada H., Umeda A., Morimura T.,
RA Uetsuka K., Doi K., Tsuji S., Kitamura D.,
RT "BASH, a novel signaling molecule preferentially expressed in B cells
RT of the bursa of Fabricius."
RL J. Immunol. 161:5804-5808(1998).
DR EMBL; AF089727; AAD12783.1; -
DR EMBL; AB015289; BAA36275.1; -
DR HSSP; P08487; 2PLD.
DR Interpro; IPR002965; P-rich_extensn.
DR Interpro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PRO1217; PRICHEXTENSN.
DR PRINTS; PRO0401; SH2DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
SQ SEQUENCE 552 AA; 61822 MW; FB232179BE38D072 CRC64;

Query Match 13.6%; Score 315.5; DB 13; Length 552;
Best Local Similarity 25.6%; Pred. No. 2,4e-17;
Matches 117; Conservative 53; Mismatches 154; Indels 133; Gaps 18;

QY 57 VPGEKCSNNDDYDPEFOLKAMPMSKILPARIQSEYADTRYFQDMMEAP---LILPKKASYST 106
DB 105 VPSEE--NPDDSYEPPESEQK-----KIPSPFISRGYIDNRTSHQOLPPIKPLPST 158
QY 107 -EAPLLPPKASVST-----ERQTRDVRMTQLEVD-----KPTFKDVSQRF 148
DB 159 PSSALPRPKKPSLPSPAKPKPLPKPRECSDENYIVPVNDNDNYIEPT----- 209
QY 149 KGFYTKINKTLPPLPPRAITL-----PKKYQLPRA--PPESSAYFAP 191
DB 210 -----ESSNP--PPAKPVNRFKPPAKSALPPPKPSIASDQOEYVEVEEELSPP 261
QY 192 KPT-FPEVQGRPRORSADFSRVLGAEES-----HQTKPPESSCPSS 233
DB 262 PVTRFTKPLPATRAQNA-EHSHMSMTRESKRLDASRNITLPLPRLNRHLPKTDHANNDE 320
QY 234 N--ONTOKS--PPAIA-----SSSYMGKHSIQARD-----HTG 263
DB 321 NHTSNTQESKEFPGAASPLPALKTSNMAVNPAPKPLPSRDFTVNEDKPTAADRRRG 380
QY 264 SMQCPACQQAASHS-----PRMLP-----YENTNSKPPPTPDEKDVQON 307
DB 381 SSHERPPLPLPSGPPKSSLOKPLVLPKVEAPSRALGTSPISSSISSTADQAGVSK 440
QY 308 WYIGEYSROAVEDYLMKNDGTFLVRDCSTKSKAEPPVLYVFGNKYVNVKIRFLESNQ 367
DB 441 AWYATATCDKRAEDALYNSKDGSLIRKSSGQDSQRYTLVVFTRKRYNIPVFIETATK 500
QY 368 QFALGTGLRGNEMFDSVEDIIEHTYTPILLIDGKD 404
DB 501 RQYALGRKKGEEYFGVAEIIRNHQHSPLVLIDSON 537

```

RESULT 10
075499 PRELIMINARY: PRT: 433 AA.
ID 075499
AC 075499;
DT 01-NOV-1998 (TREMBlrel. 08, Created)
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE B cell linker protein BLNK-S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
RT receptor activation."
RL J. Biol. Chem. 272:27362-27368(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turck C.W., Kurosaki T., Chan A.C.;
RT "BLNK: a central linker protein in B cell activation."
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20050956; PubMed=10583958;
RA Mlegishi Y., Rohrer J., Cousstan-Smith E., Lederman H.M., Pappu R.,
RA Campana D., Chan A.C., Conley M.E.;
RT "An essential role for BLNK in human B cell development."
RL Science 286:1954-1957(1999).
DR EMBL: AF068181; AAC39937.1;
DR EMBL: AF180756; AAF20383.1;
DR EMBL: AF180740; AAF20383.1; JOINED.
DR EMBL: AF180741; AAF20383.1; JOINED.
DR EMBL: AF180742; AAF20383.1; JOINED.
DR EMBL: AF180743; AAF20383.1; JOINED.
DR EMBL: AF180744; AAF20383.1; JOINED.
DR EMBL: AF180745; AAF20383.1; JOINED.
DR EMBL: AF180746; AAF20383.1; JOINED.
DR EMBL: AF180748; AAF20383.1; JOINED.
DR EMBL: AF180749; AAF20383.1; JOINED.
DR EMBL: AF180750; AAF20383.1; JOINED.
DR EMBL: AF180751; AAF20383.1; JOINED.
DR EMBL: AF180752; AAF20383.1; JOINED.
DR EMBL: AF180753; AAF20383.1; JOINED.
DR EMBL: AF180754; AAF20383.1; JOINED.
DR EMBL: AF180755; AAF20383.1; JOINED.
DR HSSP: P29354; 1BMB.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2.1.
DR ProDom: PD000093; SH2.1.
DR SMART: SM00252; SH2.1.
DR PROSITE: PS50001; SH2.1.
SQ SEQUENCE 433 AA; 48229 MW; 0B36FE9FCF5DC7DC CRC64;

Query Match 13.48; Score 309.5; DB 4; Length 433;
Best Local Similarity 25.38; Pred. NO. 5.3e-17;
Matches 99; Conservative 51; Mismatches 127; Indels 115; Gaps 12;
QY 64 NSNNYEDPEFOLTKAWSMKILPAPQSEFYADTRFQDMQMAP---LLPRAVST 120
DB 91 NADDSYEPPEPVE---QSTRVYHPLPARGETIDNRSSQ--RHSPPSKTLPSKPWPS 144
QY 121 ERQTRFDVMTQLEVDYDPTFKDVSQRKFGKYTKINKTPLPPEPRATILPKKYQPLPPA 180
DB 145 E-----KARLSTLPALTAQKRP-VPPK 167
QY 181 PP---EESSAIFAKPTPTPEVQGRPRASANDFSRVLCAGEESHQTRPSSCPSSNQN- 236

```

```

DB 168 PKGLEDEADYVP-----VEDDENYIHPTESSPEPEKGRNS 206
QY 237 ----TQSPPAIASSSYMBGKH-----SIQARDHTGSM---OHCPAORCOAA----- 276
DB 207 GAMETKSPPPAPAPPLPRAGKKPTPLKTTTPVAAQQNASSVCEKPIAPRRHRSRQGE 266
QY 277 ASHSPMLLYENTNSRPP-----TKPDEKDWQNEWTG 312
DB 267 AVQSPVPPPAOKIHKQPIPLPRFTGEGNPYDGLPSFSSNSTISEQAGVLCRPMVAG 326
QY 313 EYSROAVEDYLNKNDGTFLVDCSTKSKAPRYLVFYGNKYVNVKIFLESNOQFAL 372
DB 327 ACDRKSAEALIRNSKNDGSFLIRKSGCHSKQPYTLVYFENKRVYINIPVFIEATIKQVAL 386
QY 373 GTGLGNEMKFDVSDIEDIIEHYTFPILLIDGKD 404
DB 387 GRKKNGEYFGSVAEIIRHQHSPVLIDSON 418

RESULT 11
Q9QUN3 PRELIMINARY: PRT: 457 AA.
ID Q9QUN3
AC Q9QUN3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE SH2-containing leukocyte protein 65 (Lymphocyte antigen 57).
GN SLP-65 OR BASH OR LY57.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BALB/C; TISSUE-LYMPHOID;
RX MEDLINE=98372771; PubMed=9705962;
RA Wienands J., Schweikert J., Wollschlaed B., Junna H., Nielsen P.J.,
RA Reih M.;
RT "SLP-65: A new signalling component in B lymphocytes which requires
RT expression of the antigen receptor for phosphorylation."
RL J. Exp. Med. 188:791-795(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN-BALB/C; TISSUE-LYMPHOID;
RA Wienands J., Labollette O., Reih M.;
RT "Evidence for a preformed transducer complex organized by the B cell
RT antigen receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 938:7865-7870(1996).
RN [3]
RP SEQUENCE FROM N.A.
RX Okamoto N., Hayashi K., Tsuji S., Gotsuka R., Kitamura D.;
RT "BASH: B lymphocyte adaptor protein containing SH2 domain."
RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RX TRANSDUCER-ETN;
RA Nielsen P.J., Guenet J.L.;
RT "The murine SLP-65 gene."
RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: Y17159; CAA76666.1;
DR EMBL: AB015290; BAA34944.1;
DR EMBL: AJ298054; CAC18565.1;
DR HSSP: P23727; 1BFI.
DR MCD: MGI:96878; LY57.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2.1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2.1.
DR SMART: SM00252; SH2.1.
DR PROSITE: PS50001; SH2.1.
RW B-cell.
FT CONFLICT 133 133 S -> N (IN REF. 2).

```

```

RESULT 12
ID 088504 PRELIMINARY; PRT; 457 AA.
AC 088504;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE B cell linker protein BLNK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
RT receptor activation."
RL J. Biol. Chem. 272:27368-27368(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turck C.W., Kurosaki T., Chan A.C.;
RT "BLNK: A Central Linker Protein in B Cell Activation."
RL EMBL: AF068182; AAC40206.1;
DR HSSP: P23727; IBFI.
DR MCD: MGI:96878; Iy57.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.

```

```

RESULT 13
Q9D413
ID Q9D413 PRELIMINARY; PRT: 297 AA.
AC Q9D413;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DR 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DE 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE 4935424C13Rik protein.
GN 4935424C13Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shitagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Flieschmann W., Gaasterland T., Gissi C., King B., Koehli H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikolaev I., Pesole G., Quackenbush J.,
RA Schiml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazarrelli J., Mombaerts P.,
RA Noriote P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."

```



```
Db 201 SLDEPWTE-BEGD-----GSDHPYNSIPSKMPPPGCFIDTRLKPRPHAP-- 244
QY 184 ESSAYEAPK-----PTPE-VORGPRORSAKDFSRVLGAEEESHQTKPESSCP 231
Db 245 -DFAOFAGKEQTYQGRHLDGTFEGEDWQOTPLROGSSD---IYSTPEGKLVHAPTEGEAP 299
QY 232 SSNONTQKSPPALASSSYMFGKHSIQARDHTGSMQHCPAQRCAASHSPR-----MLPY 286
Db 300 -TYVNTQOIPR-----QANPA-AVSSAESSPRKDLFDMPF 333
QY 287 ENTNSEKP-----DPTKPDEK---DVMQNEWYIGEYSROAVEDVLMKE 326
Db 334 EDALKNQPLGPRVLSKASVECTISPVSPRAPDAKMLEELQAFETYQGEMSRKEAEGLL--- 390
QY 327 NKDGTFLVRDCSTKSKAEP--YVLVIFYGNKVYVKIRFILESNOQFALGTGLRGNEMFDS 384
Db 391 EKDDDFLVR---KSTNPGSFVLTGMHNGQAKHLLVDPE-----GTERTKDRVFD 439
QY 385 VEDIIEHY--TYPIILLIDGKDKARRKOCYLOPL 418
Db 440 ISHLINHLESPLIV-----SAGSELC-LOQPV 467
```

Search completed: April 21, 2003, 12:37:55
Job time : 33.6461 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:32:18, Search time 31.6461 Seconds

(without alignments)
2832.275 Million cell updates/sec

Title: US-09-856-061-2

Perfect score: 2316

Sequence: 1 MTSQGNKRTTKEGFDLRFQ.....QPLPLARLLTFQSSQALHE 435

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database:

SPTREMBL_21:*
1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.podent:*
12: sp.virus:*
13: sp.vertibrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriaph:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	100.0	435	11 Q9JMJ3	Q9JMJ3 mus musculus
2	2305	99.5	435	11 Q9QZJ2	Q9QZJ2 mus musculus
3	1132	48.9	376	4 Q9P2U9	Q9P2U9 homo sapien
4	410.5	17.7	533	11 Q9J2M0	Q9J2M0 mus musculus
5	410	17.7	525	11 Q9J2L0	Q9J2L0 rattus norv
6	381	16.5	530	13 Q9DG07	Q9DG07 gallus gall
7	319	13.8	456	4 O8WZ28	O8WZ28 homo sapien
8	319	13.8	456	4 O7S498	O7S498 homo sapien
9	315.5	13.6	552	13 Q9YGC1	Q9YGC1 gallus gall
10	309.5	13.4	433	4 Q75499	Q75499 homo sapien
11	297.5	12.8	457	11 Q9QUN3	Q9QUN3 mus musculus
12	287.5	12.4	457	11 Q88504	Q88504 mus musculus
13	219.5	9.5	297	11 Q9D413	Q9D413 mus musculus
14	188.5	8.1	516	11 Q54737	Q54737 mus musculus
15	167.5	7.2	471	4 Q9DCX5	Q9DCX5 homo sapien
16	165.5	7.1	594	4 Q9Z529	Q9Z529 homo sapien

17	165.5	7.1	594	4 O8TAP2	O8TAP2 homo sapien
18	150.5	6.5	486	4 Q9UCX4	Q9UCX4 homo sapien
19	150.5	6.5	812	15 O85466	O85466 y73.sarcoma
20	143.5	6.2	662	5 Q9N355	Q9N355 caenorhabdi
21	141.5	6.1	474	11 Q61120	Q61120 mus musculus
22	141	6.1	728	4 Q9UPH9	Q9UPH9 homo sapien
23	139.5	6.0	286	5 Q95PX0	Q95PX0 caenorhabdi
24	139.5	6.0	398	5 Q95PW9	Q95PW9 caenorhabdi
25	138.5	6.0	335	4 Q9GCZ8	Q9GCZ8 homo sapien
26	138.5	6.0	926	4 Q9H0K1	Q9H0K1 homo sapien
27	137.5	5.9	594	11 Q70143	Q70143 rattus norv
28	136.5	5.9	477	11 Q70142	Q70142 rattus norv
29	134	5.8	309	11 Q9QVZ0	Q9QVZ0 mus sp. shb
30	134	5.8	461	11 Q63789	Q63789 rattus norv
31	134	5.8	1599	5 Q18892	Q18892 caenorhabdi
32	133.5	5.8	217	11 Q9CX99	Q9CX99 mus musculus
33	132.5	5.7	904	5 Q9VMJ4	Q9VMJ4 drosophila
34	132.5	5.7	944	5 Q8R9B3	Q8R9B3 drosophila
35	130.5	5.6	722	13 Q8UUR2	Q8UUR2 xenopus lae
36	129	5.6	559	11 Q91Z52	Q91Z52 mus musculus
37	129	5.6	645	6 Q95JY5	Q95JY5 macaca fasc
38	128.5	5.5	112	11 P70305	P70305 mus musculus
39	128.5	5.5	393	2 Q93CA3	Q93CA3 treponema p
40	128.5	5.5	854	11 Q9QZS8	Q9QZS8 mus musculus
41	128.5	5.5	1430	11 Q8VHK2	Q8VHK2 rattus norv
42	128	5.5	623	6 Q95J17	Q95J17 macaca fasc
43	128	5.5	651	11 P97504	P97504 mus musculus
44	128	5.5	1160	11 Q9ESC8	Q9ESC8 mus musculus
45	128	5.5	4498	13 Q93291	Q93291 fugu rubrip

ALIGNMENTS

RESULT 1	PRELIMINARY:	PRT:	435 AA.
Q9JMJ3			
AC Q9JMJ3:			
DT 01-OCT-2000 (TREMBLrel. 15, Created)			
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)			
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)			
DE MIST.			
OS CLINK OR MIST.			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
OX NCBI_TaxID=10090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RX MEDLINE-20208989; PubMed-10744659;			
RA Goitsuka R., Kanazashi H., Sasanna H., Fujimura Y., Hida Y.,			
RA Tatsuono A., Ra C., Hayashi K., Kitamura D.,			
RT "A BASH/SLP-76-related adaptor protein MIST/Clink involved in Ige			
RT receptor-mediated mast cell degranulation.";			
RL Int. Immunol. 12:573-580(2000).			
DR EMBL: AB021220; BAA96240.1; -			
DR HSSP: P23727; 2PNB.			
DR MGD; MGI:1351468; Clink.			
DR InterPro; IPR000980; SH2.			
DR Pfam; PF00017; SH2; 1.			
DR PRINTS; PR00401; SH2DOMAIN.			
DR ProDom; PD000093; SH2; 1.			
DR SMART; SM00252; SH2; 1.			
DR PROSITE; PS50001; SH2; 1.			
SQ SEQUENCE 435 AA; 49513 MW; 4D1BD3E2F0C61ED6 CRC64;			
Query Match	100.0%;	Score 2316;	DB 11; Length 435;
Best Local Similarity	100.0%;	Pred. No. 8.3e-180;	
Matches 435;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY 1 MTSQGNKRTTKEGFDLRFQNSVSLKNRSPVLSAKKRCRAVLEPLPHRRRLACVPGG 60			
DB 1 MTSQGNKRTTKEGFDLRFQNSVSLKNRSPVLSAKKRCRAVLEPLPHRRRLACVPGG 60			


```

OY 61 EKCSNNDEDEPFOELKAMPMSKILPARPIQSEYADTRFQDMMEAPILLPPKASVST 120
    |||||
Db 61 EKCSNNDEDEPFOELKAMPMSKILPARPIQSEYADTRFQDMMEAPILLPPKASVST 120
OY 121 ERQTRDVRMTQLEEVDPKPTFKDVRSGRFGKFKYTKINKTLPPLPPRAITLPPKKYQPLPPA 180
    |||||
Db 121 ERQTRDVRMTQLEEVDPKPTFKDVRSGRFGKFKYTKINKTLPPLPPRAITLPPKKYQPLPPA 180
OY 181 PPESSAYFAKPTPEVQGRPRORSADKFSRYLGAEEESHQTKPESSCPSSNONTQKS 240
    |||||
Db 181 PPESSAYFAKPTPEVQGRPRORSADKFSRYLGAEEESHQTKPESSCPSSNONTQKS 240
OY 241 PPAIASSYMPGKHSIOADHTGSMOCHPAORCOAASHSPMLPYENTNSEKPDPTKPD 300
    |||||
Db 241 PPAIASSYMPGKHSIOADHTGSMOCHPAORCOAASHSPMLPYENTNSEKPDPTKPD 300
OY 301 EKDVQWQEWYIGEYSRQAVEDVLMKENKDGTFVLRDCSTKSKAEPYLVYVFGNKYVNVK 360
    |||||
Db 301 EKDVQWQEWYIGEYSRQAVEDVLMKENKDGTFVLRDCSTKSKAEPYLVYVFGNKYVNVK 360
OY 361 IRFLESNOQFALGTGLRGNMEDSVEDIIEHTYTFPILLIDGDKAARRKQCYLTQPLPL 420
    |||||
Db 361 IRFLESNOQFALGTGLRGNMEDSVEDIIEHTYTFPILLIDGDKAARRKQCYLTQPLPL 420
OY 421 ARLLLTQYSSQALHE 435
    |||||
Db 421 ARLLLTQYSSQALHE 435

```

RESULT 2
09QZ2 PRELIMINARY: PRT: 435 AA.

```

ID 09QZ2 01-MAY-2000 (TREMBlrel. 13, Created)
AC 09QZ2 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBlrel. 20, Last annotation update)
DE CLNK.
GN CLNK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6;
RX MEDLINE=20029811; PubMed=10562326;
RA Cao M.Y., Davidson D., Yu J., Latour S., Veillelette A.;
RT "Clunk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
    stimulated Hemopoietic Cells."
RL J. Exp. Med. 190:1537-1534(1999).
DR EMBL; AF187819; AAF14299.1;
DR HSSP; P23727; 2PNB.
DR MGD; MGI:1351468; Clnk.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
SO SEQUENCE 435 AA; 49492 MW; 5CD27EC971FC0EAS CRC64;

```

Query Match 99.5%; Score 2305; DB 11; Length 435;
Best Local Similarity 99.5%; Pred. No. 6.5e-179;
Matches 433; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

OY 1 MTSQGNKRTTEGFGDLRFQNVSLIKNSWPSLSASAKRCRAVLEPLDHRNLAGVPGG 60
    |||||
Db 1 MTSQGNKRTTEGFGDLRFQNVSLIKNSWPSLSASAKRCRAVLEPLDHRNLAGVPGG 60
OY 61 EKCSNNDEDEPFOELKAMPMSKILPARPIQSEYADTRFQDMMEAPILLPPKASVST 120
    |||||
Db 61 EKCSNNDEDEPFOELKAMPMSKILPARPIQSEYADTRFQDMMEAPILLPPKASVST 120
OY 139 TFKDVNSQRKFKYKINKTLPPLPPRAITLPPKKYQPLPPAPPESSAYFAKPTPEV 198
    |||||

```

```

OY 121 ERQTRDVRMTQLEEVDPKPTFKDVRSGRFGKFKYTKINKTLPPLPPRAITLPPKKYQPLPPA 180
    |||||
Db 121 ERQTRDVRMTQLEEVDPKPTFKDVRSGRFGKFKYTKINKTLPPLPPRAITLPPKKYQPLPPA 180
OY 181 PPESSAYFAKPTPEVQGRPRORSADKFSRYLGAEEESHQTKPESSCPSSNONTQKS 24
    |||||
Db 181 PPESSAYFAKPTPEVQGRPRORSADKFSRYLGAEEESHQTKPESSCPSSNONTQKS 2
OY 241 PPAIASSYMPGKHSIOADHTGSMOCHPAORCOAASHSPMLPYENTNSEKPDPTKPD 300
    |||||
Db 241 PPAIASSYMPGKHSIOADHTGSMOCHPAORCOAASHSPMLPYENTNSEKPDPTKPD 300
OY 301 EKDVQWQEWYIGEYSRQAVEDVLMKENKDGTFVLRDCSTKSKAEPYLVYVFGNKYVNVK 360
    |||||
Db 301 EKDVQWQEWYIGEYSRQAVEDVLMKENKDGTFVLRDCSTKSKAEPYLVYVFGNKYVNVK 360
OY 361 IRFLESNOQFALGTGLRGNMEDSVEDIIEHTYTFPILLIDGDKAARRKQCYLTQPLPL 420
    |||||
Db 361 IRFLESNOQFALGTGLRGNMEDSVEDIIEHTYTFPILLIDGDKAARRKQCYLTQPLPL 420
OY 421 ARLLLTQYSSQALHE 435
    |||||
Db 421 ARLLLTQYSSQALHE 435

```

RESULT 3
09P209 PRELIMINARY: PRT: 376 AA.

```

ID 09P209 01-OCT-2000 (TREMBlrel. 15, Created)
AC 09P209 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE MIST (Fragment).
GN MIST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20208989; PubMed=10744659;
RA Goitsuka A., Kanazashi H., Sasanna H., Fujimura Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clunk involved in Ige
    receptor-mediated mast cell degranulation."
RL Int. Immunol. 12:573-580(2000).
DR EMBL; AB032369; BAA96241.1;
DR HSSP; P23727; 2PNB.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000093; SH2; 1.
DR SMART; SM00252; SH2; 1.
DR PROSITE; PSS0001; SH2; 1.
FT NON_TER 1
FT NON_TER 376
SO SEQUENCE 376 AA; 43542 MW; 389421B629B028B2 CRC64;

```

Query Match 48.9%; Score 1132; DB 4; Length 376;
Best Local Similarity 60.9%; Pred. No. 8.2e-84;
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

```

OY 19 FQNVSLIKNSWPSLSASAKRCRAVLEPLDHRNLAGVGGKNSNNDEDEPFOELK 78
    |||||
Db 1 FQNVSLIKNSWPSLSASAKRCRAVLEPLDHRNLAGVGGKNSNNDEDEPFOELK 78
OY 79 AWPMSKILPARPIQSEYADTRFQDMMEAPILLPPKASVSTQTRDVRMTQLEEVDPK 138
    |||||
Db 79 AWPMSKILPARPIQSEYADTRFQDMMEAPILLPPKASVSTQTRDVRMTQLEEVDPK 138
OY 61 TWQSIKILPARPIKESYADTHFKYAMDPPLDTRTSSISQPTWNTQ-TRLEVDKP 119
    |||||

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:32:18 ; Search time 27.3539 Seconds

(without alignments)
2832.275 Million cell updates/sec

Title: US-09-856-061-4

Perfect score: 2020
Sequence: 1 FQNSLPKRNKSPRINSATG.....GLRGDEKPDSDIEDIIEHYKN 376

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

SPREMBL 21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	2020	100.0	376	4	Q9P2U9
2	1132	56.0	435	11	Q9JMJ3
3	1132	56.0	435	11	Q9QZJ2
4	334	16.5	552	13	Q9YGC1
5	317.5	15.7	530	13	Q9DG07
6	316.5	15.7	525	11	Q920L0
7	314.5	15.6	533	11	Q922M0
8	289.5	14.3	433	4	Q75499
9	289	14.3	456	4	Q8WV28
10	282	14.0	457	11	Q9QUN3
11	273	13.5	457	11	Q88504
12	177	8.8	297	11	Q9DA13
13	177	8.8	297	11	Q9DA13
14	153	7.6	516	11	Q54737
15	147	7.3	724	11	P70304
16	144.5	7.2	1616	4	O15054

17	142	7.0	722	13	Q8UWJ2	Q8UWJ2 xenopus lae
18	138.5	6.9	981	15	Q92809	Q92809 abelson mur.
19	135	6.7	559	11	Q91252	Q91252 mus. musculu
20	134.5	6.7	662	5	Q9N355	Q9N355 caenorhabdi
21	132	6.5	112	11	P70305	P70305 mus musculu
22	132	6.5	764	5	Q24708	Q24708 drosophila
23	131.5	6.5	471	4	Q9UCX5	Q9UCX5 homo sapien
24	131.5	6.5	485	3	Q9UTP6	Q9UTP6 schizosacch
25	130.5	6.5	548	5	Q9VPU1	Q9VPU1 drosophila
26	129.5	6.4	594	4	Q92529	Q92529 homo sapien
27	129.5	6.4	594	4	Q8RAP2	Q8RAP2 homo sapien
28	128	6.3	17352	5	Q95YM2	Q95YM2 procanbarus
29	126.5	6.3	728	4	Q9UPH3	Q9UPH3 homo sapien
30	126	6.2	377	11	Q55032	Q55032 mus musculu
31	126	6.2	377	13	P79956	P79956 xenopus lae
32	126	6.2	461	11	Q63789	Q63789 ratius norv
33	125.5	6.2	533	11	Q8VEN5	Q8VEN5 mus musculu
34	125.5	6.2	1258	11	Q8R457	Q8R457 mus musculu
35	125	6.2	269	11	Q8CT18	Q8CT18 mus musculu
36	125	6.2	377	11	Q92279	Q92279 mus musculu
37	125	6.2	377	11	Q99M51	Q99M51 mus musculu
38	125	6.2	460	4	Q96HD7	Q96HD7 homo sapien
39	125	6.2	595	13	Q92124	Q92124 xenopus lae
40	125	6.2	597	11	Q64509	Q64509 mus musculu
41	125	6.2	823	6	Q9TTY2	Q9TTY2 canis famli
42	124.5	6.2	217	11	Q9CX99	Q9CX99 mus musculu
43	124	6.1	410	5	Q24218	Q24218 drosophila
44	123.5	6.1	863	4	Q9UPV4	Q9UPV4 homo sapien
45	123	6.1	927	4	Q9UEG4	Q9UEG4 homo sapien

ALIGNMENTS

RESULT 1

Q9P2U9 PRELIMINARY: PRT: 376 AA.

Q9P2U9: Q9P2U9: 01-OCT-2000 (TReMBLrel. 15, Created)

Q9P2U9: Q9P2U9: 01-OCT-2000 (TReMBLrel. 15, Last sequence update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

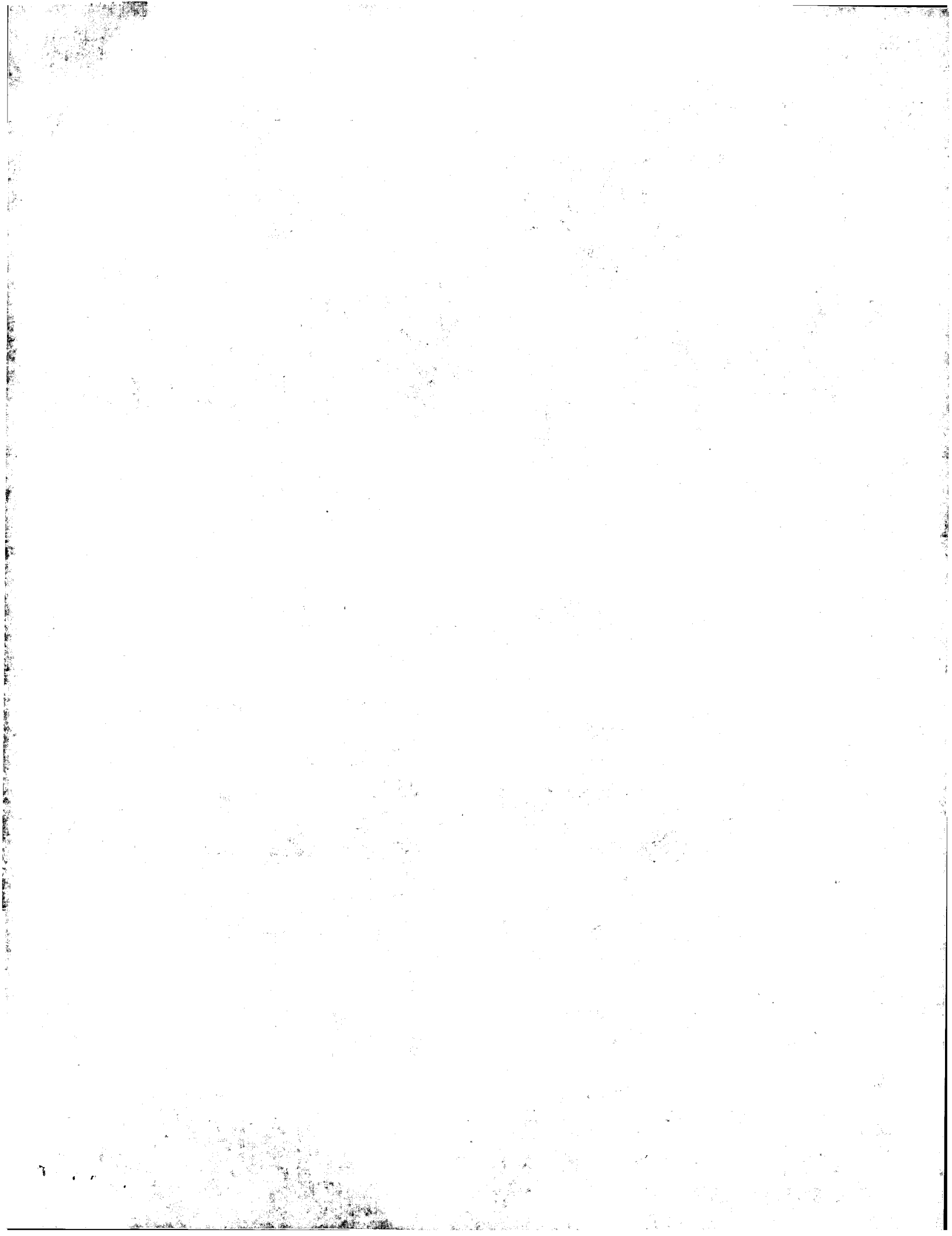
Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)

Q9P2U9: Q9P2U9: 01-JUN-2002 (TReMBLrel. 21, Last annotation update)



```

Db 1 FONFSLPKNRSPRINSATGOYORANKPLLDWERNFAVLGAKGSHDDYDDPELMEE 60
QY 61 TWOSIKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTLEVDKPI 120
Db 61 TWOSIKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTLEVDKPI 120
QY 121 SMDVASONIKGDASVKKKKIPLPPLPPLITLTKKYQPLPEPESSRPLSGRHTPEVOG 180
Db 121 SMDVASONIKGDASVKKKKIPLPPLPPLITLTKKYQPLPEPESSRPLSGRHTPEVOG 180
QY 181 MFSQSLNDLSEVLAEKVPHNOKRPESTHLENTQOEIPLAISSSFTTSHNSVQND 240
Db 181 MFSQSLNDLSEVLAEKVPHNOKRPESTHLENTQOEIPLAISSSFTTSHNSVQND 240
QY 241 HRGMOPCSPORCOPPASCSPHENILPYKYSWRRPFPKRSRDKVDQHNEMVIGEYSROA 300
Db 241 HRGMOPCSPORCOPPASCSPHENILPYKYSWRRPFPKRSRDKVDQHNEMVIGEYSROA 300
QY 301 VEEAFMKENKDGSLVRCSTSKSEEPYVLAFFENKYNVYKIRFLERNOQFALGTGLRG 360
Db 301 VEEAFMKENKDGSLVRCSTSKSEEPYVLAFFENKYNVYKIRFLERNOQFALGTGLRG 360
QY 361 DEKEDSVEDIIEHYKN 376
Db 361 DEKEDSVEDIIEHYKN 376

```

RESULT 2

```

Q9JMJ3 PRELIMINARY; PRT; 435 AA.
ID 09JMJ3
AC 09JMJ3: 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE MIST.
GN CLNK OR MIST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20208989; PubMed=10744655;
RA Gotsuka R., Kanazashi H., Sasakuma H., Fujimura Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clnk involved in Ige
RT receptor-mediated mast cell degranulation.";
RT Int. Immunol. 12:573-580(2000).
DR EMBL: AB021220; BAA96240.1; -.
DR HSSP: P23727; 2PNB.
DR MGD: MGI:1351468; Clnk.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2.1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 435 AA; 49513 MW; 4D1BD3E2FC061ED6 CRC64;

```

Query Match 56.0%; Score 1132; DB 11; Length 435;
 Best Local Similarity 60.9%; Pred. No. 3.2e-85;
 Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

```

QY 1 FONFSLPKNRSPRINSATGOYORANKPLLDWERNFAVLGAKGSHDDYDDPELMEE 60
Db 19 FONVSLKNSWPSLSAKRCRAVLEPLDHRNRNLGAVGGEKCSNNNDYEDPEFQLK 78
QY 61 TWOSIKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTLEVDKPI 119
Db 79 AWMSPKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTLEVDKPI 138
QY 120 ISDVRSONIKGDASVKKKKIPLPPLPPLITLTKKYQPLPEPESSRPLSGRHTPEV 178

```

```

Db 139 TEKDVRSQRFKGFYKTIKNTPLPEPPPAITLPPKKYQPLPEPESSASVAPKPTPEV 198
QY 179 QGMFSQSLRLDSEVLAEKVPHNOKRPESTHLENTQOEIPLAISSSFTTSHNSVQND 238
Db 199 QGPRORSKADPSRYLGAEEESHQOTPESSCPSSNONTQKSPALAIASSSYMPGKHSIOA 258
QY 239 RDHFGMOPCSPORCOPPASCSPHENILPYKYSWRRPFPKRSRDKVDQHNEMVIGEYSR 298
Db 259 RDHFGMOCHCAQOQAASHSP--RMPLPENTNSEKPDPTKPEKDVQWQNEVIGEYSR 316
QY 299 QAVEAFMKENKDGSLVRCSTSKSEEPYVLAFFENKYNVYKIRFLERNOQFALGTGL 358
Db 317 QAVEDVLAKENKDGSTPLVRCSTSKSEEPYVLAFFENKYNVYKIRFLERNOQFALGTGL 376
QY 359 RGDEKEDSVEDIIEHY 374
Db 377 RGNEMFDSVEDIIEHY 392

```

RESULT 3

```

Q9QZE2 PRELIMINARY; PRT; 435 AA.
ID 09QZE2
AC 09QZE2: 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE CLNK.
GN CLNK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6;
RX MEDLINE=20029811; PubMed=10562326;
RA Cao M.Y., Davidson D., Yu J., Latour S., Veillette A.;
RT "Clnk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
RT stimulated Hemopoietic Cells.";
RT J. Exp. Med. 190:1527-1534(1999).
DR EMBL: AF187819; AAF14299.1; -.
DR HSSP: P23727; 2PNB.
DR MGD: MGI:1351468; Clnk.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2.1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 435 AA; 49492 MW; 5CD27EC971F0CBA5 CRC64;

```

Query Match 56.0%; Score 1132; DB 11; Length 435;
 Best Local Similarity 60.9%; Pred. No. 3.2e-85;
 Matches 229; Conservative 41; Mismatches 102; Indels 4; Gaps 3;

```

QY 1 FONFSLPKNRSPRINSATGOYORANKPLLDWERNFAVLGAKGSHDDYDDPELMEE 60
Db 19 FONVSLKNSWPSLSAKRCRAVLEPLDHRNRNLGAVGGEKCSNNNDYEDPEFQLK 78
QY 61 TWOSIKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTLEVDKPI 119
Db 79 AWMSPKILPAPRIKESSEYADTHYFKVAMDTPLDTRTSISIGOPTWNTOTLEVDKPI 138
QY 120 ISDVRSONIKGDASVKKKKIPLPPLPPLITLTKKYQPLPEPESSRPLSGRHTPEV 178
Db 139 TEKDVRSQRFKGFYKTIKNTPLPEPPPAITLPPKKYQPLPEPESSASVAPKPTPEV 198
QY 179 QGMFSQSLRLDSEVLAEKVPHNOKRPESTHLENTQOEIPLAISSSFTTSHNSVQND 238
Db 199 QGPRORSKADPSRYLGAEEESHQOTPESSCPSSNONTQKSPALAIASSSYMPGKHSIOA 258
QY 239 RDHFGMOPCSPORCOPPASCSPHENILPYKYSWRRPFPKRSRDKVDQHNEMVIGEYSR 298

```


GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:32:18 ; Search time 27.3539 Seconds
(without alignments)
2832.275 Million cell updates/sec

Title: US-09-856-061-4
Perfect score: 2020
Sequence: 1 FQNFSLPKRNSWPRINSATG.....GLNGDEKFDSDIEDIEHYKN 376

Scoring table: BIOSUM62
Gapop 10.0, Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_minc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rudent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2020	100.0	376	4	Q9P2U9
2	1132	56.0	435	11	Q9JMJ3
3	1132	56.0	435	11	Q9QZC2
4	334	16.5	552	13	Q9YGC1
5	317.5	15.7	530	13	Q9DG07
6	316.5	15.7	525	11	Q92010
7	314.5	15.6	533	11	Q922M0
8	289.5	14.3	433	4	Q75499
9	289	14.3	456	4	Q8WV28
10	289	14.3	456	4	Q75498
11	282	13.5	457	11	Q9QUN3
12	273	13.5	457	11	Q88504
13	177	8.8	297	11	Q9D413
14	153	7.6	516	11	Q54737
15	147	7.3	724	11	P70304
16	144.5	7.2	1616	4	O15054

17	142	7.0	722	13	Q8UWU2	Q8UWU2 xenopus lae
18	138.5	6.9	981	15	Q92809	Q92809 abelson mur
19	135	6.7	559	11	Q91252	Q91252 mus musculus
20	134.5	6.7	662	5	Q9N3S5	Q9N3S5 caenorhabdit
21	132	6.5	112	11	P70305	P70305 mus musculus
22	132	6.5	764	5	Q24708	Q24708 drosophila
23	131.5	6.5	471	4	Q9UCX5	Q9UCX5 homo sapien
24	131.5	6.5	485	3	Q9UPP6	Q9UPP6 schizosacch
25	130.5	6.5	548	5	Q9YPU1	Q9YPU1 drosophila
26	129.5	6.4	594	4	Q92529	Q92529 homo sapien
27	129.5	6.4	594	4	Q8RAP2	Q8RAP2 homo sapien
28	128	6.3	17352	5	Q95YV2	Q95YV2 procambarus
29	126.5	6.3	728	4	Q9UPH9	Q9UPH9 homo sapien
30	126	6.2	377	11	Q55032	Q55032 mus musculus
31	126	6.2	377	13	P79956	P79956 xenopus lae
32	126	6.2	461	11	Q83789	Q83789 ratius norv
33	125.5	6.2	533	11	Q8EVN5	Q8EVN5 mus musculus
34	125.5	6.2	1258	11	Q8R457	Q8R457 mus musculus
35	125	6.2	269	11	Q9CT18	Q9CT18 mus musculus
36	125	6.2	377	11	Q92279	Q92279 mus musculus
37	125	6.2	377	11	Q99M51	Q99M51 mus musculus
38	125	6.2	460	4	Q96HD7	Q96HD7 homo sapien
39	125	6.2	595	13	Q92124	Q92124 xenopus lae
40	125	6.2	597	11	Q64509	Q64509 mus musculus
41	125	6.2	823	6	Q9TTY2	Q9TTY2 canis famli
42	124.5	6.2	217	11	Q9CX99	Q9CX99 mus musculus
43	124	6.1	410	5	Q24218	Q24218 drosophila
44	123.5	6.1	863	4	Q9UPV4	Q9UPV4 homo sapien
45	123	6.1	927	4	Q9UDE4	Q9UDE4 homo sapien

ALIGNMENTS

RESULT 1

Q9P2U9 ID Q9P2U9 PRELIMINARY; PRT; 376 AA.
AC Q9P2U9;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MIST (Fragment).
GN MIST.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=20208989; PubMed=10744659;
RX Gotsuka R., Kanazashi H., Sasamura H., Fujimura Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clnk involved in Ige
receptor-mediated mast cell degranulation.";
RL Int. Immunol. 12:573-580(2000).
DR EMBL: AB032369; BAA96241.1; .
DR HSSP: P23727; 2PMB.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2.1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2.1.
DR SMART: SM00252; SH2.1.
DR PROSITE: PS50001; SH2.1.
FT NON_TER 1
FT NON_TER 376
SQ SEQUENCE 376 AA; 43542 MW; 389421B629B028B2 CRC64;

Query Match 100.0%; Score 2020; DB 4; Length 376;
Best Local Similarity 100.0%; Pred. No. 1.6e-158;
Matches 376; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 FQNFSLPKRNSWPRINSATGQYORNMKPLLDERNFVAALDGAKSHSDDDYDPELRMBE 60
|||||

```

Db      1 FONSLPKNRSPRINSATGOYQRMNKPILLDMERNFAVLGAKGSHDDYDDPELMBE 60
Qy      61 TWOSIKILPARIKESEYADTHFKVAMDTPLDTRTISIGOPTWNTOTLEEVDPRI 120
Db      61 TWOSIKILPARIKESEYADTHFKVAMDTPLDTRTISIGOPTWNTOTLEEVDPRI 120
Qy      121 SRDVRSONIKGASVAKNKIPLPPRPPLTLTKKYOPLPEPESSRPLSQHHTPEVQ 180
Db      121 SRDVRSONIKGASVAKNKIPLPPRPPLTLTKKYOPLPEPESSRPLSQHHTPEVQ 180
Qy      181 MFSQISLRDLSEVLEAEKYPHNRKPESTHLENONTOEIPLAISSSFTSNHVSQVRD 240
Db      181 MFSQISLRDLSEVLEAEKYPHNRKPESTHLENONTOEIPLAISSSFTSNHVSQVRD 240
Qy      241 HGGMOQPCSPQRCOPPCASCPHENILPYKYTSMRPPFPKSRDKRDQVQHNEMYGESYRQA 300
Db      241 HGGMOQPCSPQRCOPPCASCPHENILPYKYTSMRPPFPKSRDKRDQVQHNEMYGESYRQA 300
Qy      301 VEEAFMKENKDGSLVRCSTFKSKEPEYVLAIFYENKYNKIRPLENOOPALGTGLRG 360
Db      301 VEEAFMKENKDGSLVRCSTFKSKEPEYVLAIFYENKYNKIRPLENOOPALGTGLRG 360
Qy      361 DEKEDSVEDIIEHYKN 376
Db      361 DEKEDSVEDIIEHYKN 376

```

RESULT 2

```

09JMJ3      PRELIMINARY;      PRT;      435 AA.
AC      09JMJ3;
DT      01-OCT-2000 (TREMBLrel. 15, Created)
DT      01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT      01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE      MIST.
GN      CLNK OR MIST.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_Taxid=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=20208989; PubMed=10744659;
RA      Goltisuka R., Kanazashi H., Sasannuna H., Fujimura Y., Hidaka Y.,
RA      Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT      "A BASH/SLP-76-related adaptor protein MIST/Clnk involved in Ige
RT      receptor-mediated mast cell degranulation.";
RL      Int. Immunol. 12:573-580(2000).
DR      EMBL: AB021220; BAA96240.1; -.
DR      HSSP: P23727; 2PNB.
DR      MGD: MGI:1351468; Clnk.
DR      InterPro: IPR000980; SH2.
DR      Pfam: PF00017; SH2; 1.
DR      PRINTS: PR00401; SH2DOMAIN.
DR      ProDom: PD00093; SH2; 1.
DR      SMART: SM00252; SH2; 1.
DR      PROSITE: PS50001; SH2; 1.
SQ      SEQUENCE 435 AA; 49513 MW; 4D1BD3E2F0C61ED6 CRC64;

```

Query Match Best Local Similarity 56.0%; Score 1132; DB 11; Length 435; Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

```

Qy      1 FONSLPKNRSPRINSATGOYQRMNKPILLDMERNFAVLGAKGSHDDYDDPELMBE 60
Db      19 FONSLPKNRSPRINSATGOYQRMNKPILLDMERNFAVLGAKGSHDDYDDPELMBE 60
Qy      61 TWOSIKILPARIKESEYADTHFKVAMDTPLDTRTISIGOPTWNTOTLEEVDPRI 119
Db      79 AMPSMKILPARIKESEYADTHFKVAMDTPLDTRTISIGOPTWNTOTLEEVDPRI 119
Qy      120 SRDVRSONIKGASVAKNKIPLPPRPPLTLTKKYOPLPEPESSRPLSQHHTPEV 178
Db      120 SRDVRSONIKGASVAKNKIPLPPRPPLTLTKKYOPLPEPESSRPLSQHHTPEV 178

```

```

Db      139 TFKDVRSQRFKFKYTKINKTLPPLPPRAITLPPKKYQPLPPAPPESSAYFAKPTPEV 198
Qy      179 QGMPQSILRDLSEVLEAEKYPHNRKPESTHLENONTOEIPLAISSSFTSNHVSQVRD 238
Db      199 QGMPQSILRDLSEVLEAEKYPHNRKPESTHLENONTOEIPLAISSSFTSNHVSQVRD 238
Qy      239 RDHGGMOQPCSPQRCOPPCASCPHENILPYKYTSMRPPFPKSRDKRDQVQHNEMYGESYR 298
Db      239 RDHGGMOQPCSPQRCOPPCASCPHENILPYKYTSMRPPFPKSRDKRDQVQHNEMYGESYR 298
Qy      299 QAVEAFMKENKDGSLVRCSTFKSKEPEYVLAIFYENKYNKIRPLENOOPALGTGL 358
Db      317 QAVEAFMKENKDGSLVRCSTFKSKEPEYVLAIFYENKYNKIRPLENOOPALGTGL 358
Qy      359 RGDEKEDSVEDIIEHY 374
Db      377 RGDEKEDSVEDIIEHY 392

```

RESULT 3

```

09QZE2      PRELIMINARY;      PRT;      435 AA.
AC      09QZE2;
DT      01-MAY-2000 (TREMBLrel. 13, Created)
DT      01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT      01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE      CLNK.
GN      CLNK.
OS      Mus musculus (Mouse).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX      NCBI_Taxid=10090;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      STRAIN=C57BL/6;
RX      MEDLINE=20029811; PubMed=10562326;
RA      Cao M.Y., Davidson D., Yu J., Latour S., Veillette A.;
RT      "Clnk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
RT      stimulated Hemopoietic Cells ";
RL      J. Exp. Med. 190:1527-1534(1999).
DR      EMBL: AF187819; AAF14299.1; -.
DR      HSSP: P23727; 2PNB.
DR      MGD: MGI:1351468; Clnk.
DR      InterPro: IPR000980; SH2.
DR      Pfam: PF00017; SH2; 1.
DR      PRINTS: PR00401; SH2DOMAIN.
DR      ProDom: PD00093; SH2; 1.
DR      SMART: SM00252; SH2; 1.
DR      PROSITE: PS50001; SH2; 1.
SQ      SEQUENCE 435 AA; 49492 MW; 5CD27EC971FC0EAS CRC64;

```

Query Match Best Local Similarity 56.0%; Score 1132; DB 11; Length 435; Matches 229; Conservative 41; Mismatches 102; Indels 4; Gaps 3;

```

Qy      1 FONSLPKNRSPRINSATGOYQRMNKPILLDMERNFAVLGAKGSHDDYDDPELMBE 60
Db      19 FONSLPKNRSPRINSATGOYQRMNKPILLDMERNFAVLGAKGSHDDYDDPELMBE 60
Qy      61 TWOSIKILPARIKESEYADTHFKVAMDTPLDTRTISIGOPTWNTOTLEEVDPRI 119
Db      79 AMPSMKILPARIKESEYADTHFKVAMDTPLDTRTISIGOPTWNTOTLEEVDPRI 119
Qy      120 SRDVRSONIKGASVAKNKIPLPPRPPLTLTKKYOPLPEPESSRPLSQHHTPEV 138
Db      120 SRDVRSONIKGASVAKNKIPLPPRPPLTLTKKYOPLPEPESSRPLSQHHTPEV 138
Qy      139 TFKDVRSQRFKFKYTKINKTLPPLPPRAITLPPKKYQPLPPAPPESSAYFAKPTPEV 198
Db      179 QGMPQSILRDLSEVLEAEKYPHNRKPESTHLENONTOEIPLAISSSFTSNHVSQVRD 238
Qy      199 QGMPQSILRDLSEVLEAEKYPHNRKPESTHLENONTOEIPLAISSSFTSNHVSQVRD 238
Db      239 RDHGGMOQPCSPQRCOPPCASCPHENILPYKYTSMRPPFPKSRDKRDQVQHNEMYGESYR 298
Qy      239 RDHGGMOQPCSPQRCOPPCASCPHENILPYKYTSMRPPFPKSRDKRDQVQHNEMYGESYR 298

```


Db 259 RDHTSGMOHCPAORCAASHSP--RMLPYENTNSEKPDPTKPDDEKDVQWQNEWYIGEXSR 316
 QY 299 QAVEBAFKENKDGSEFLVRCDSYTSKEEPPYLAIFYENKYNVKTRETERNOQFALGTGL 358
 Db 317 QAVEBVLKKNKDGSEFLVRCDSYTSKAPYLYVYFNGKNYVNVKTRFLESNOQFALGTGL 376
 QY 359 RGDEKDSVEDITEHY 374
 Db 377 RGNEMFDSVEDITEHY 392

RESULT 4

QYGC1 PRELIMINARY: PRT: 552 AA.
 ID QYGC1
 AC QYGC1
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE B cell linker protein BLNK.
 GN BLNK OR BASH.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99146381; PubMed-10023776;
 RA Ishai M., Kurosaki M., Pappu R., Okawa K., Ronko I., Fu C.,
 RA Shibata M., Iwamatsu A., Chan A.C., Kurosaki T.;
 RT "BLNK required for coupling syk to PLC gamma 2 and Rac1-JNK in B
 cells";
 RL Immunity 10:117-125(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-99049801; PubMed-9834055;
 RA Gotsuka R., Fujimura Y., Mamada H., Umeda A., Morimura T.,
 RA Uetsuka K., Doi K., Tsuji S., Kitamura D.;
 RT "BASH, a novel signaling molecule preferentially expressed in B cells
 of the bursa of Fabricius";
 RL J. Immunol. 161:5804-5808(1998).
 RL EMBL; AF089727; AAD12783.1; -;
 DR EMBL; AB015289; BAA36275.1; -;
 DR HSSP; P08487; 2PID.
 DR InterPro; IPR002965; P_Rich_extensn.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PRO1217; PRICEXTENSN.
 DR PRODOM; PD00401; SH2DOMAIN.
 DR PRODOM; PD00093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS0001; SH2; 1.
 SQ SEQUENCE 552 AA; 61822 MW; FB232179BE3BD072 CRC64;

Query Match 16.5%; Score 334; DB 13; Length 552;
 Best Local Similarity 24.3%; Pred. No. 2, 9e-19;
 Matches 112; Conservative 60; Mismatches 157; Indels 132; Gaps 15;

QY 32 WEERFAVLIDGAKGHS-----DDYDDELMEETWOSIKILPARPIESES 78
 Db 83 WSDDFSDENPDGHSDEMYVSEENPDSDYEPPESEK-----KIPSPISRGEX 138
 QY 79 AD---HYKRVAMDPPLDTRTSI-----SIGOPTWMTQTL-----ERDKPI 120
 Db 139 ADKRISHQPLPNKPLSPSSALPRPKKPSLPSPAAKPKLPLKRECSDEEDNTIVPY 198
 QY 121 SRDVSQNIKGASV-----RKNKIPLPRLPILPKYQPLPPEESS 165
 Db 199 DNDNDYIETPESSTPPAKPYNRMKPPAKGALPTPPKPLASMOQEVYEPPEEEL 258
 QY 166 RPPLSORHTPEYQMP-----QISLDLSEVLEAEK-----VPHNQKPESTHL 211
 Db 259 SPPPVTRFKP---LPATRAQNAEHSNHSMTRESPKLDASPNILPLPNRILHPXTDHE 314

QY 212 LEN-----QNTQE-----IPLAISSS-----FTSNHSYON 238
 Db 315 ANNDENHSFNTQSKPPPGGAAPSLPRALKTSNAVPAKPCPLPSDPTTVEDKFTA 374
 QY 239 RDHRGMO---PCSPORCOPPAS-----GSPHENILPYKYSWRP 275
 Db 375 ADRRGSSHEFFLPPLPGTPEKSSLOKPLVLPKVEAPSRALGTPSHSISSTA--- 431
 QY 276 PPKRSRDKVOHNEWYIGEXSROAVEBAFKENKDGSEFLVRCDSYTSKEEPPYLAIFYE 335
 Db 432 ----DODAGYASKAMVATCDKRYAEDALYRSNKGSEFLIRKSSGDSROPYTLVYFN 486
 QY 336 NKVYVVKIRPLERNOQFALGTGLRGDEKDEKDSVEDITEHYKN 376
 Db 487 RRVNIPIRFESTROYALGRKCGEERPDVAELVENHQH 527

RESULT 5

QYGC07 PRELIMINARY: PRT: 530 AA.
 ID QYGC07
 AC QYGC07
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE SLP-76 adaptor protein.
 GN SLP76.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-20451093; PubMed-10939315;
 RA Ishai M., Kurosaki M., Inabe K., Chan A.C., Suganuma K., Kurosaki T.;
 RT "Involvement of LAT, Gads, and Grb2 in Complementatation of SLP-76 to
 the Plasma Membrane";
 RL J. Exp. Med. 192:847-856(2000).
 DR EMBL; AF226986; AAG18493.1; -;
 DR HSSP; P12931; 1SHD.
 DR InterPro; IPR001660; SAM.
 DR InterPro; IPR000980; SH2.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PRO0401; SH2DOMAIN.
 DR PRODOM; PD00093; SH2; 1.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS0001; SH2; 1.
 SQ SEQUENCE 530 AA; 59580 MW; CF6CBCEBD79909A CRC64;

Query Match 15.7%; Score 317.5; DB 13; Length 530;
 Best Local Similarity 26.9%; Pred. No. 6, 2e-18;
 Matches 105; Conservative 53; Mismatches 145; Indels 87; Gaps 14;

QY 46 HSDDYDDELMEETWOSIKILPARPI-KESFYAD-----THYFKYAMD----- 89
 Db 140 HSDGYEPPPSNNDENAHNV-IFPAKSLANNNDYIDRPTSSSHQPPVPPQRPGPSAP 198
 QY 90 -----TPLPLDTRTSISIGP-----TWNTQRLERVDKPISRDVSQNIKG 131
 Db 199 ASFGGRGASLPAPFPLPGNNDRVKSKPPAPSIDRSTKPLDRLGPPERE---NPVPG 255
 QY 132 DASVRKKNKILPLPPR---PLILPKYQPLPPEPSSR---PLLSQHTPEYQMGMS 183
 Db 256 RKGHEPEKLTLPOLRALGSQLAMPK-----PVPPSDRYERGNPSPKQILPVKQGMQ 310
 QY 184 QISLRDSEVLEAEKYPHNOQKPESTHLENON-----QOELPLAISSSFTSNHS 235
 Db 311 QKRPEEEDHIIPORAVPQISLPYSSNTPPSKIKAPKPGSGSIFGASARLSMS---- 367
 QY 236 VONRDRGMOQSPORCOPPASCPHENILPYKYTSWRPPFD-----KSRDRKD 285

Db 368 -----SLPRPFLGN--SRSPSRGT-----ADLRPLIPSRQTAHOTNTEDEED 412

OY 286 VOHNEMYIGESROAVEAEAFMKENKDGSLVDCSTKKEEYVLAAYENKVVVKIRF 345

Db 413 SLNDEMVAAYISRPAAEAALRKINODGTFVLVDSRSKRTTHRYLMLVLYKDVYNIQIRY 472

OY 346 LERNOQFALGTGLRGDEKFDVSEDIIEHYK 375

Db 473 QEQDQTYLLGTGLKGEDEFFSSVAHIIDYEQ 502

RESULT 6

O920L0 PRELIMINARY: PRT: 525 AA.

AC O920L0:

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE SLP-76 adaptor protein.

GN SLP76.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE FROM N.A.

RA Fujii Y., Goltchuka G., Wakahara S., Hara T., Nakao T.;

RT "Molecular cloning of SLP-76, a 76-kDa adaptor protein in RBL-2H3.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: AB072980; BAB71779.1; -

DR InterPro: IPR000980; SH2.

DR Pfam: PF00017; SH2; 1.

DR PRODOM: PD000093; SH2; 1.

DR PROSITE: PSS0001; SH2; 1.

SQ SEQUENCE 525 AA; 59521 MW; 0B1CDD40DABE7372 CRC64;

Query Match 15.78; Score 316.5; DB 11; Length 525;

Best Local Similarity 25.98; Pred. No. 7,4e-18;

Matches 117; Conservative 49; Mismatches 112; Indels 173; Gaps 19;

OY 20 GOYORMNKPLDWERNFAAVLDGAKGSHDDYDDELMEETWQSIILPARPIKESEYA 79

Db 126 GDYSPN-----DEEELVDDAA-----DYEPSPNDEALON-SILPRKPR----- 166

OY 80 DTHYFKVAMDTPLDTRTISISIGOPTWNTQTRLEVDKPRISRDVRSQNTKGDASVKNK 139

Db 167 -----PMTNSK-----YIDRPTGKVSQP----- 186

OY 140 IPLRPRLITLPKKYQRLP-----PEP-----ESSRPLSQ-- 171

Db 187 -PVPQRPMAALP---PLPTGRNHSNHEPSPNNRNHKTAKLRAPSIDRSTKPRDLRSL 241

OY 172 -----NHTPEVQ--GMPQSILRDLSE--VLEAK--VPH 201

Db 242 APRLDREPLIGKPKSPKAPLGRNHLPKIQKRLPRAMORHENEIGLTLAKRPVPR 301

OY 202 NQKRP-----STHLENQNTQ-----EPLAISSSSF 229

Db 302 HGRPRDRONDEDDVHQRLPHPSLRPSMSNTFSPKTPSPKNTFPLPHMPCALSESNI 361

OY 230 TTSNHSQVNDHRCGMOPCS--PORCOPRASCSPHEN--ILPYKYTSMRPPPKRSDK 284

Db 362 -----GFQOSASLPSYFSOGSPSSRPVRNRYVNLRLPVPN-RPQPSGEE 407

OY 285 DVQHNEMYIGESROAVEAEAFMKENKDGSLVDCSTKKEEYVLAAYENKVVVKIR 344

Db 408 SPLDEEYVYSITRPEAEALRKINODGTFVLVDSRSKRTYNNPYLVLYLTKDVYNIQIR 467

OY 345 FLERNOQFALGTGLRGDEKFDVSEDIIEHYK 375

Db 468 YQESQVYLLGTGLKGEDEFLVSIIIDYFR 498

RESULT 7

O922M0 PRELIMINARY: PRT: 533 AA.

AC O922M0:

DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)

DE Lymphocyte cytosolic protein 2.

GN LCP2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RA Strausberg R.;

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL: BC006948; AAH06948.1; -

DR MGD: MGI:1321402; LCP2.

DR InterPro: IPR000980; SH2.

DR Pfam: PF00017; SH2; 1.

DR PRODOM: PD000093; SH2; 1.

DR PROSITE: PSS0001; SH2; 1.

SQ SEQUENCE 533 AA; 60238 MW; 50AEA025EF0DAD01 CRC64;

Query Match 15.68; Score 314.5; DB 11; Length 533;

Best Local Similarity 30.08; Pred. No. 1.1e-17;

Matches 121; Conservative 44; Mismatches 169; Indels 69; Gaps 18;

OY 20 GOYORMNKPLDWERNFAAVLDGAKGSHDDYDDELMEETWQSIILPARPI--KESE 77

Db 126 GDYSPN-----DEEELVDDAA-----DYEPSPNNEALOS-SILPNSFHTNSM 172

OY 78 YAD-THYFKVAMDTPL-DLTRTISISIGOPTWNTQTRLEVDKPRISRDVRSQNTKGDASV 135

Db 173 YIDRPTGKVSQRPVPRPLPKRPLP-PLPTGRNHSPLSPHNPHEPSSGNNK----- 226

OY 136 RKNKIRP-----PRPRLITLPKKYQRLP-----PEPSSRPLSQRTPEVQ--GM 181

Db 227 -TAKLRAPSIDRSTKPRPLDSLAPLDREPTLIGKPKSPKAPLGRNHLRKIQKRL 284

OY 182 PQSILRDLSEVL-----EAEKYPHNRKPESTHLE--NQNTQEIPLAISSSSFSTSN 233

Db 285 PRAMDREHNERELRGVTRKRPVPRHRCGRPRENDEDDVHQRLPQSLPSMSNTFPS 344

OY 234 HSYQ-----NRDHRGMPCS--PORCOPRASCSP--HENILPYKITS 272

Db 345 RSYQPSKNTFPLAHMPCAFSESNTIGFQOSASLPSYFSOGPNRPLRSEGRNLPLRVPN 404

OY 273 WRPFPRKSDKDVQHNEMYIGESROAVEAEAFMKENKDGSLVDCSTKKEEYVLAAY 332

Db 405 -RPQPSGEEPTRLDEEWVYSITRPEAEALRKINODGTFVLVDSRSKRTYNNPYLVLY 463

OY 333 FYENKYYNKAIRLERNOQFALGTGLRGDEKFDVSEDIIEHYK 375

Db 464 LYKDKVYNIQIRYQESQVYLLGTGLKGEDEFLVSIIIDYFR 506

RESULT 8

O75499 PRELIMINARY: PRT: 433 AA.

AC O75499:

DT 01-NOV-1998 (TREMBLrel. 08, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)

DE B cell linker protein BLNK-S.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

Db	391	NGEEYGSVAETIRNHQ	408
Db	391	NGEEYGSVAETIRNHQ	408
RESULT 9			
Q8WV28	PRELIMINARY;	PRT;	456 AA.
AC	Q8WV28;		
DT	01-MAR-2002 (TREMBLrel. 20, Created)		
DT	01-MAR-2002 (TREMBLrel. 20, Last sequence update)		
DT	01-JUN-2002 (TREMBLrel. 21, Last annotation update)		
DT	B-cell linker.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=TONSIL;		
RA	Strausberg R.;		
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC018906; AAH18906.1; ..		
DR	InterPro; IPR000980; SH2.		
DR	Pfam; PF00017; SH2; 1.		
DR	ProDom; PD000093; SH2; 1.		
DR	SMART; SM00252; SH2; 1.		
DR	PROSITE; PS50001; SH2; 1.		
SO	SEQUENCE 456 AA; 50465 MW; ED6D424A035D1792 CRC64;		
Query Match	14.3%; Score 289; DB 4; Length 456;		
Best Local Similarity	25.6%; Pred. No. 1,1e-15;		
Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;			
QY	43 AKGSHDDDDYDDELMEETWOSIKILPARPIKESEADTHYKVMADPPLDTRTSISI 102		
DB	88 AEENADDSYEPPVEQ-----TRYPALPFAARGEY-----IDNRSS--- 125		
QY	103 GQPTNTQRLERVDKPLSRDVRSONIKDASVKKKILPLPPRPRLTLKKYOLPPREP 162		
DB	126 -----ORHSPFSKTLPSK-----PSWPSEKARLTSLPLALQRPQ-VPPRP 168		
QY	163 -----ESSRP-----PLSORHTFPPEVQ-----GMPQSITSLD 189		
DB	169 KLLDEADLVVYPVEDNDENYIHPRESSPPPEKAMVNRKSNSTPASPPTASGRN 228		
QY	190 LSEVLEAEK-----VPHNQRP-----ESTHLENTQETPLAISSFTTSNHSVQ 237		
DB	229 -SGAWETKSPPPAAPSPLPRAGKKPTPLKTPVVASQN-----ASSVCEEKPIP 277		
QY	238 NNDHGG--MQCSPRCRCPASCSHENILRY-KT-SWRP-----PFRKRDGRKVOYN 289		
DB	278 AERHRRSSHROAVALOSPPVPPAKOIHQRPILPRFTTEGNGPLVDDPLRFSFNSNTISQ 337		
QY	290 E-----HYIEGYSQAAVEAEAFMKENKGGSFLLRDCSTFKSEEPYLVAFYENKYVNRK 342		
DB	338 EAGVCLCKPYAGACDRKSADEALHRSNKKGCSFLIRKSSGHDQKPYTLVVFENKRYNIP 397		
QY	343 IRFLERNQOFALGTGLRGDEKFDVSDIETIHYKN 376		
DB	398 VRIETATKOYALGRKKNGEYGSVAETIRNHQ 431		
RESULT 10			
ID	075498		
AC	075498;		
DT	01-NOV-1998 (TREMBLrel. 08, Created)		
DT	01-NOV-1998 (TREMBLrel. 08, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)		
DE	B cell linker protein BLNK.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
RT receptor activation."
RN J. Biol. Chem. 272:27362-27368(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turk C.W., Kurosaki T., Chan A.C.;
RT "BLNK: a central linker protein in B cell activation."
RN Immunity 9:93-103(1998).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20050936; PubMed=10583958;
RA Muegelsht Y., Rohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
RA Campana D., Chan A.C., Conley M.E.;
RT "An essential role for BLNK in human B cell development."
RN Science 286:1954-1957(1999).
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98068180; AAC39936.1;
DR EMBL: AF180756; AAF20382.1; -
DR EMBL: AF180740; AAF20382.1; JOINED.
DR EMBL: AF180741; AAF20382.1; JOINED.
DR EMBL: AF180742; AAF20382.1; JOINED.
DR EMBL: AF180743; AAF20382.1; JOINED.
DR EMBL: AF180744; AAF20382.1; JOINED.
DR EMBL: AF180745; AAF20382.1; JOINED.
DR EMBL: AF180746; AAF20382.1; JOINED.
DR EMBL: AF180747; AAF20382.1; JOINED.
DR EMBL: AF180748; AAF20382.1; JOINED.
DR EMBL: AF180749; AAF20382.1; JOINED.
DR EMBL: AF180750; AAF20382.1; JOINED.
DR EMBL: AF180751; AAF20382.1; JOINED.
DR EMBL: AF180752; AAF20382.1; JOINED.
DR EMBL: AF180753; AAF20382.1; JOINED.
DR EMBL: AF180754; AAF20382.1; JOINED.
DR EMBL: AF180755; AAF20382.1; JOINED.
DR HSSP: P29354; 1BM8.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 456 AA; 50466 MW; 95F1D5485D03D397 CRC64;

Query Match 14.3%; Score 289; DB 4; Length 456;
Best Local Similarity 25.6%; Pred. No. 1,1e-15;
Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;

QY 43 AKGSHDDYDPELMETWOSIKILPARPIKESYADTHYFKVAMDTPLDPTRTSISI 102
DB 88 AENADDSYEPPEVE-----TRPVHPLPARGFY-----IDNRSS--- 135
QY 103 GQPTWNTQTLEVDKRTISDVNSQNTKGASVRKKNIPPLPPPLITLTKKXQPLPEP 162
DB 126 -----QRHSPFSKTLPSK-----PSWPESEKARLTSTLPAALAQKQ-VPPKP 168
QY 163 -----ESSRP-----PLSQHRTPEVQ-----GMSQSILRD 189
DB 169 KGLLEDEADYVVEVDENDEYIHTPESSSPPEKAPVNRNSTPNSSTPASPCTAGRN 228
QY 190 LSEVLEAEK-----VPHNORP-----ESTHLEENQNTQETPLAISSSEFTSNHSVQ 237
DB 229 -SGAMETKSPPPAPSPPLPRAGKKPTPLKTTTPVAVSQ-----ASSVCEKPIR 277
QY 238 NRDRHG--WQPSQRCQPCPASCSPHENLPI-KYT-SWRP-----PPKRSRDKVQNH 289
DB 278 AERHRRSSHRQEAIVQSPVPPAPQKQIHOKPIPLPRFTEGNGPTVDGDLPSFSNSTISQ 337
QY 290 E-----WYIGETSRQAVAEAFMKENKDGSLVYRDCSTKSKPEPVAVFENKYYNKK 342

```

```

DB 338 EAGVLCRWYAGACDRSAEALHRNKGDSFLLRKSSGHDSPYTLVFFENKRYNIP 397
QY 343 IRFLERNOQFALGTGLRGDEKEDSVEDIIEHYKN 376
DB 398 VRFEATKQVALGKRKKNGEEFSSVAELIRNHQ 431

RESULT 11
Q9QUN3 PRELIMINARY; PRT; 457 AA.
ID Q9QUN3
AC Q9QUN3;
DT 01-MAY-2000 (Trembl, 13, Created)
DT 01-MAY-2000 (Trembl, 13, Last sequence update)
DT 01-MAR-2002 (Trembl, 20, Last annotation update)
DE SH2-containing leukocyte protein 65 (lymphocyte antigen 57).
GN SLP-65 OR BASH OR LY57.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=LYMPHOID;
RX MEDLINE=98372771; PubMed=9705962;
RA Wiemands J., Schweikert J., Wollischled B., Jumaa H., Nielsen P.J.,
RA Rech M.;
RT "SLP-65: A new signalling component in B lymphocytes which requires
RT expression of the antigen receptor for phosphorylation."
RN J. Exp. Med. 188:791-795(1998).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/C; TISSUE=LYMPHOID;
RA Wiemands J., Larbolette O., Rech M.;
RT "Evidence for a preformed transducer complex organized by the B cell
RT antigen receptor."
RN Proc. Natl. Acad. Sci. U.S.A. 93:7865-7870(1996).
RN [3]
RP SEQUENCE FROM N.A.
RA Okamoto N., Hayashi K., Tsuji S., Gotsuka R., Kitamura D.;
RT "BASH: B lymphocyte adaptor protein containing SH2 domain."
RN Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC TRANSDOMAIN-ETN;
RA Nielsen P.J., Guenet J.L.;
RT "The murine SLP-65 gene."
RN Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: Y17159; CAA76666.1; -
DR EMBL: AB015290; BAA34944.1; -
DR EMBL: AJ298054; CAC18565.1; -
DR HSSP: P23727; 1BFT.
DR MGD: MGT:96876; LY57.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
KW B-cell.
FT CONFLICT 133 133 S -> N (IN REF. 2).
FT CONFLICT 147 147 A -> T (IN REF. 2).
FT CONFLICT 148 150 RLA -> GIG (IN REF. 2).
FT CONFLICT 197 198 PP -> AT (IN REF. 2).
FT CONFLICT 391 391 N -> K (IN REF. 2).
FT CONFLICT 444 446 TRD -> SKH (IN REF. 2).
SQ SEQUENCE 457 AA; 50670 MW; 66C93D4FDDF9D260 CRC64;

Query Match 14.0%; Score 282; DB 11; Length 457;
Best Local Similarity 23.8%; Pred. No. 4,3e-15;
Matches 96; Conservative 58; Mismatches 157; Indels 92; Gaps 15;

QY 32 WERNFAAVLDGAKGHS-----DDDYDPELMETWOSIKILPARPIKESY 78

```

```

DB 64 MSDDSDSYENDEHSEDEMYVMAPEETGDDSYEPPEAQ-----QTRVHPALPPTRGEX 119
OY 79 ADHYFKYAMDPLPLDTRTSISIGOPTWNTOTRLERVDKPLSRDVRSONIKGDA5VRKN 138
DB 120 VDNR-----SSQHSPPFSKTLPS--KPSW-----PSAKARLASTLPAPNSLOKP 162
OY 139 KIPPLPPRLI-----TLPKYOPLPPEP-----ESSRPLSQRHPEP 176
DB 163 QVP-FKPKDLDEADYVVPVEDNDENYIHPRESSPPPEAKAPMVRSTKPPSSSKHNSP 221
OY 177 --EVOGMPQSILRLSEVLEAEKYPHNQKRESETHLENTOTLEIPLAISSSFTTSMH 234
DB 222 PGTVAGRNGSVWDSKSSLPAPSPPLRAGKRP-ATPL-----KTTVP-PLPANSVCEK 275
OY 235 SVQNDHRCG--MOPCSFORCOPPASCSPHENILPYKTSWMPREPK-----279
DB 276 PYPAERHRSSHRDPTVQSPVFPPTQKPVHQKRVPL-----PRPEAGSPAADGPFHSF 329
OY 280 -----RSRKRQVQHNEMWIGEYSRQAVEAEAFMKENKDSFLVRDOSTKSKEEPPYLAVF 333
DB 330 PENSTFPADQEAELGKPMWAGACDRKSAEALHRSNKDGSFLIRKSSGHDSKQPYTLVAF 389
OY 334 YENKYVNVKIRFLERNOQFALGTGLRGDEKFDSDVEDIIEHYKN 376
DB 390 FKKRYVNIPIRFEATKOYALGKKNGEYFGSVVEIVNSHOH 432

RESULT 12
O88504 PRELIMINARY: PRT: 457 AA.
AC 088504:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE B cell linker protein BLNK.
GN LY57.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN 11]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Gbd2, and Yav after B cell antigen
RT receptor activation."
RT J. Biol. Chem. 272:27362-27368(1997).
RN 12]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turk C.W., Kurosaki T., Chan A.C.;
RT "BLNK: A Central Linker Protein in B Cell Activation."
RT Immunity 9:93-103(1998).
DR EMBL: AF068182; AAC40206.1;
DR HSSP: P23372; 1BFT.
DR MGD: MGI:96878; LY57.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 457 AA; 50803 MW; 66D235796A6C45F0 CRC64;

Query Match 13.5%; Score 273; DB 11; Length 457;
Best Local Similarity 23.8%; Pred. No. 2,4e-14;
Matches 96; Conservative 60; Mismatches 155; Indels 92; Gaps 16;

OY 32 WENFAAVLDGAKGHS-----DDDDDDDELMEETWQSIIKILPAPRKSEY 78
DB 64 WSDFDSDYENDEHSEDEMYVMAPEETGDDSYEPPEAQ-----QTRVHPALPPTRGEX 119

```

```

OY 79 ADHYFKYAMDPLPLDTRTSISIGOPTWNTOTRLERVDKPLSRDVRSONIKGDA5VRKN 138
DB 120 VDNR-----SSQHSPPFSKTLPS--KPSW-----PSAKARLASTLPAPNSLOKP 162
OY 139 KIPPLPPRLI-----TLPKYOPLPPEP-----ESSRPLSQRHPEP 176
DB 163 QVP-FKPKDLDEADYVVPVEDNDENYIHPRESSPPPEAKAPMVRSTKPPSSSKHNSP 221
OY 177 --EVOGMPQSILRLSEVLEAEKYPHNQKRESETHLENTOTLEIPLAISSSFTTSMH 234
DB 222 PGTVAGRNGSVWDSKSSLPAPSPPLRAGKRP-ATPL-----KTTVP-PLPANSVCEK 275
OY 235 SVQNDHRCG--MOPCSFORCOPPASCSPHENILPYKTSWMPREPK-----280
DB 276 PYPAERHRSSHRDPTVQSPVFPPTQKPVHQKRVPL-----PRPEAGSPAADGPFHSF 329
OY 281 -----SDRK-DVQHNEMWIGEYSRQAVEAEAFMKENKDSFLVRDOSTKSKEEPPYLAVF 333
DB 330 PENLTFPADQEGELGKPMWAGACDRKFAEALHRSNKDGSFLIRKSSFGHDSKQPYTLVAF 389
OY 334 YENKYVNVKIRFLERNOQFALGTGLRGDEKFDSDVEDIIEHYKN 376
DB 390 FKKRYVNIPIRFEATKOYALGKKNGEYFGSVVEIVNSHOH 432

RESULT 13
O9D413 PRELIMINARY: PRT: 297 AA.
AC 09D413:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE 4933424C13RIK.
GN 4933424C13RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN 11]
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamagata I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita K., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection."
RL Nature 409:685-690(2001).
DR EMBL: AK016886; BAB30480.1;
DR HSSP: P16277; 1BTK.
DR MGD: MGI:1918380; 4933424C13RIK.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SQ SEQUENCE 297 AA; 32015 MW; BECE6E8ACF70685C CRC64;

```

```

Query Match      8.8%; Score 177; DB 11; Length 297;
Best Local Similarity 25.6%; Pred. No. 1,1e-06;
Matches 68; Conservative 32; Mismatches 106; Indels 60; Gaps 9;

QY 138 NRIPLPPPLTLTKKYPPLPEPPSSRPPLSORHTPEVOGMSQ-----184
DB 37 SKPPLEPPQSTMA-----RGLPINSFPPIRPTSGYH-FPLKATAMPQAPKQGFVGRGR 90
QY 185 -ISLR-----DLSEVLEAEKVPNHNKRPSTHLENNQOETPLAISS-----SF 229
DB 91 GTSARAVTKKPPEDITYLEEPPDPVPLTRSLSSKALI-----PPVPLPTSGLPKSVAGY 145
QY 230 TTSNHSVONDRHGRGQPCSPQPCPPASCSPHENILPKYKTSWBPPEKRSRDKRDVOHN 289
DB 146 QEARCAGMDGALKAGRR-LSASSIAPALSSVAEN-----GSLIGQ 185
QY 290 EWTIGEYSQAVEEAFPMKENKDGSPFLVRDCSPKSKKEEPPVLAFFENKYNKIFLERN 349
DB 186 PWSGNCROSOVERALLHFKQKDGATKVRLLSSGPHSSQPTTLAVLLRGVFNIPRIQLOGG 245
QY 350 QQFALG-TGLRGDEKFDVSDIIEHY 374
DB 246 HHYALGREGRNHEELFSSVAAWQHY 271

RESULT 14
ID 054737 PRELIMINARY; PRT; 516 AA.
AC 054737;
DT 01-JUN-1998 (TREMBLrel. 06, Created)
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE BCA protein (fragment).
GN Lys7 OR BCA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAIB/C;
RA MEDLINE=98346794; PubMed=9683264;
RA Gangi-Peterson L., Peterson S., Shapiro L., Golding A., Caricchio R.,
RA Cohen D.I., Margulies D.H., Cohen P.L.;
RL "Bca -- An Activation-related B-cell Gene.";
DR EMBL: AJ222814; CAAL1002.1; -.
DR MGD: MGI:96878; Lys7.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2. 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
DR NON_TER 1
SQ SEQUENCE 516 AA; 57822 MW; F31E65A7DC676FDD CRC64;

Query Match      7.6%; Score 153; DB 11; Length 516;
Best Local Similarity 20.3%; Pred. No. 0.00022;
Matches 77; Conservative 46; Mismatches 130; Indels 126; Gaps 15;

QY 32 WERNFAVLDAKGS-----DDYDPELRMEETQSIKILPARIKSEY 78
DB 98 WSDPFDSDYEPDHSOSEKVMPEETGDSYEPPEAQ-----QTRVHPALPFTRGEX 153
QY 79 ADHYFVAMDTPLDTRTISISIGPTWMTQTRLEKVRPISR--DVRSQNIKGDASVR 136
DB 154 -----VNRSS-----QTLRPSAHFVSPAGLQKRAKGLG 184
QY 137 KNTKPLP-----PPRLITLTKKYQPLRP-----EPSSRPPLSQ-----172
DB 185 ASTLPANSLQKQVPPKPKROLLEADYVVPVEDNDENYIHPRESSASALLRLRLPMSID 244

```

```

QY 173 ---HTFPEV-----QGMPSQISL-----RDISEVLEAEKVP-----HN 202
DB 245 QPSQTVPOSTCASNVDAGQVGSSTPILCLPHNPHYGLGSLDHLRLPLFLPYRMHQ 304
QY 203 QNRKPESTHLE-NQNTQETPLAISSSFTTSHSVONDRHGRGQPCSPORCQPPASCSP 261
DB 305 MEVKKSLFLLSANRRSSIRHODTVQSPVPPYQKPYHOK-----PVLPRKAGSPAAQCP 357
QY 262 HENILPYRTYSWRPPFPKRSRDKDVQ---HNWTIGEYSQAVEEAFPMKENKDGSPFLVR 317
DB 358 FHS-----FPNSTFADQDEALLGKPMYVAVPNRKSASEALHRSNKDGSFLIR 405
QY 318 DCSTKSKKEEPPVLAFFEN 336
DB 406 KSSGHSQPTTLVAFFKS 424

RESULT 15
ID P70304 PRELIMINARY; PRT; 724 AA.
AC P70304;
DT 01-FEB-1997 (TREMBLrel. 02, Created)
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Phosphoinositide 3-kinase p85alpha.
GN PIK3R1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN(5); PART OF CLONE;
RA MEDLINE=97079666; PubMed=8921377;
RA Friedman D.A., Cantley L.C., Carpenter C.L.;
RT "Structural organization and alternative splicing of the murine
RL phosphoinositide 3-kinase p85 alpha gene.";
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
DR EMBL: U50413; AAC52847.1; -.
DR HSSP: P23727; 1BFI.
DR MGD: MGI:97583; PIK3R1.
DR InterPro: IPR001720; PI3Kinase_P85.
DR InterPro: IPR00198; RhogAP.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00620; RhogAP; 1.
DR Pfam: PF00017; SH2; 2.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PR00678; PI3KINASEP85.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD000093; SH2; 2.
DR SMART: SM00324; RhogAP; 1.
DR SMART: SM00252; SH2; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50002; SH3; 1.
DR KINASE: SH3 domain.
SQ SEQUENCE 724 AA; 83501 MW; BCC21211DD46F1C17 CRC64;

Query Match      7.3%; Score 147; DB 11; Length 724;
Best Local Similarity 21.7%; Pred. No. 0.0011;
Matches 76; Conservative 38; Mismatches 97; Indels 140; Gaps 17;

QY 141 PLPPRPLTLTKKYPPLPEPPSSR-----PLSQRHTPEVOGMSQISLR 188
DB 85 PTPKPRP---PR---PLVAPGSSKTEADTEQALPLPLAQCFAPPDV-APPLLIKLL 136
QY 189 DLSEV-LEAEKVPNHNKRPSTHLE-----NQNTQETPLAISSSFTTSHSVONR 239
DB 137 EATEKKGLESTLYKTOSSNPALRLDLCDAASVDLEMTDVHVLADAKRYLADIPN- 195
QY 240 DHRGMOQPCSP-----ORCOPRASC-----SPHENILPYKY-----270

```

```

Db 196 -----PVIPIVAVYNNEMMSLAQELQSPEDCLOLKKLIRLPNIPHOQWLTLOYLKHEF 248
Qy 271 -----TSW-----273
Db 249 KLSQASSKNLNNARVLEIFSPVLFEPFAASSDNTEHLIKATEILISTEMNEROPAPALP 308
Qy 274 -RPPF-----KSDRKVOQHNEWYIGEXYRQAVEAFMKENKOSFLVRDOSTKSEE 326
Db 309 PKPKFTTVANNMNNMSLODAEWYWDISREYNEK-LRDTADGTFLVRDASTKMGD 367
Qy 327 PYVLAIFYENKYYNVK-IRFLERNQOFALGTGIRGDEKFDSEDIIEHYKN 376
Db 368 -YTLT---PRKGGNNKLIRKIFHRDCKYGFSDPL---TNSYVELINHYRN 410

```

Search completed: April 21, 2003, 12:37:57
 Job time : 29.3539 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:15:31 ; Search time 73.6821 Seconds

(without alignments)
6314.347 Million cell updates/sec

Title: US-09-856-061-3

Perfect score: 2020

Sequence: 1 ttccagactcagctgcgc.....catcgactactacaagatt 1129

Scoring table:

BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US09856061/runat_21042003_113019_3744/app_query.fasta_1.3150
-DB=SPREMBL_21 -OFMT=fastan -SUFFIX=n2p -rspt -MINMATCH=0.1 -LOOPL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09856061.qcgn.1.1_324_@runat_21042003_113019_3744 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MAP -LARGEOUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

SPREMBL_21.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_minc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_verticbrate.*
14: sp_unclassified.*
15: sp_rvirus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2020	100.0	376	Q9P2U9	Q9p2u9 homo sapien

2	1132	56.0	435	11	Q9JMJ3	Q9jMJ3 mus musculus
3	1132	56.0	435	11	Q9QZK2	Q9qZK2 mus musculus
4	334	16.5	552	13	Q9YGC1	Q9yGC1 gallus gall
5	318.5	15.8	530	13	Q9DC07	Q9dC07 gallus gall
6	316.5	15.7	525	11	Q920L0	Q920L0 rattus norv
7	314.5	15.6	533	11	Q922M0	Q922M0 mus musculus
8	291	14.4	433	4	Q75499	Q75499 homo sapien
9	290.5	14.4	456	4	Q75498	Q75498 homo sapien
10	289	14.3	456	4	Q8WV28	Q8wV28 homo sapien
11	282	14.0	457	11	Q9QUN3	Q9qUN3 mus musculus
12	273	13.5	457	11	Q88504	Q88504 mus musculus
13	177	8.8	297	11	Q9D413	Q9d413 mus musculus
14	153	7.6	516	11	Q54737	Q54737 mus musculus
15	147	7.3	724	11	P70304	P70304 mus musculus
16	144.5	7.2	1616	4	Q15054	Q15054 homo sapien
17	142	7.0	722	13	Q8UUN2	Q8uUN2 xenopus lae
18	138.5	6.9	981	15	Q92809	Q92809 abelson mur
19	135	6.7	559	11	Q91252	Q91252 mus musculus
20	134.5	6.7	662	5	Q9N355	Q9n355 caenorhabdi
21	132	6.5	112	11	P70305	P70305 mus musculus
22	132	6.5	764	5	Q24708	Q24708 drosophila
23	131.5	6.5	471	4	Q9UCX5	Q9uCX5 homo sapien
24	131.5	6.5	485	3	Q9UTP6	Q9uTP6 schizosacch
25	130.5	6.5	548	5	Q9VPU1	Q9vPU1 drosophila
26	129.5	6.4	594	4	Q92529	Q92529 homo sapien
27	129.5	6.4	594	4	Q8TAP2	Q8tAP2 homo sapien
28	128	6.3	17352	5	Q95YM2	Q95YM2 procabarus
29	127.5	6.3	728	4	Q9UPH9	Q9uPH9 homo sapien
30	127	6.3	927	11	Q9UEG4	Q9UEG4 homo sapien
31	126	6.2	377	11	Q55032	Q55032 mus musculus
32	126	6.2	377	13	P79956	P79956 xenopus lae
33	126	6.2	461	11	Q63789	Q63789 rattus norv
34	125.5	6.2	533	11	Q8VEN5	Q8vEN5 mus musculus
35	125.5	6.2	823	6	Q9TVY2	Q9tVY2 canis famli
36	125.5	6.2	1258	11	Q8R457	Q8r457 mus musculus
37	125	6.2	269	11	Q9CT18	Q9cT18 mus musculus
38	125	6.2	377	11	Q92279	Q92279 mus musculus
39	125	6.2	377	11	Q99M51	Q99M51 mus musculus
40	125	6.2	460	4	Q96HD7	Q96HD7 homo sapien
41	125	6.2	595	13	Q92124	Q92124 xenopus lae
42	125	6.2	597	11	Q64509	Q64509 mus musculus
43	124.5	6.2	217	11	Q9CX99	Q9cX99 mus musculus
44	124.5	6.2	2592	3	Q9P3J0	Q9P3J0 neurospora
45	124	6.1	410	5	Q24218	Q24218 drosophila

ALIGNMENTS

RESULT 1
ID Q9P2U9 PRELIMINARY: PRT: 376 AA.
AC Q9P2U9
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE MIST (Fragment).
GN MIST.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP
SEQUENCE FROM N.A.
RX MEDLINE:20208989; PubMed:10744659;
RA Gotsuka R., Kanazashi H., Sasanuma H., Fujimura Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.,
RT "A BASH/SLP-76-related adaptor protein MST/Clnk involved in Ige
RT receptor-mediated mast cell degranulation.";
RT Int. Immunol. 12:573-580(2000).
DR EMBL: AB032369; BAA96241.1; -;
DR HSSP: P23727; 2PNB.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.

Db	59	GLYGLYGLULYSYSKSNBSNBSRNBSNBSNRYGLIUSAPROGLIUPHGLINLEULENYS	78
OY	181	ACATGGCAGTCGATTAATAATTATTCACGCCGAGCCTATTAAGGAATCTGAATATGCAGAT	240
Db	79	AlatirProserMetLysIleuProAlaArgProIleGlnGlnSerGlyIthrLysIleAsp	98
OY	241	ACACACATTTTCAAGGTGCATAGGACACTCCCTTCCTTCGTAGACACAGACGCTTCATC	300
Db	99	ThirArgIyrPheGlnAspMetMetGluAlaProLeuLeuLeuProPolysLalaserVal	118
OY	301	TCCATTGGACGCCGACCTGGACACACAG---ACGAGTTTGGAAAGGTGGACAAACCC	357
Db	119	SerThirGluArgGlnIthrArgAspValArgMetThirGlnLeuGlnGluValAspLysPro	138
OY	358	ATTTCACGGGAGCCTCAGAAAGCCCAAAACATTAAAGAGATGCATCCGTAGAAGAAACAAG	417
Db	139	ThirPheLysAspValArgSerGlnArgPheLysGlyPheLysTyrThirLysIleAsnLys	158
OY	418	ATTTCCTTACCACCTCTCTGGCCTCTCATTAACCTTCCGAAAGATACCACCCCTTGGCC	477
Db	159	ThirProLeuProProProAlaIleThrLeuProLysLysTyrIleProLeuPro	178
OY	478	CGT---GAGCGGGAGACACACAGGCCACCGCTTATTCACAGACACACCTTTCACAGATGC	534
Db	179	ProAlaProProGlnGlnSerSerAlaTyrPheAlaProLysProThirPheProGluVal	198
OY	535	CAGGAATGCCACGATCAGATTAAGGCACTTAAGTAGAGGCTCCTTGACACAGAAAA	594
Db	199	GlnArgGlyProArgGlnArgSerIleAlaLysAspPheSerArgValLeuGlyIleGlnGlu	218
OY	595	GTTCCCTATAAACACAGAGGAAGCCTGAATCACTCATCTGTGTAAAGAACCAAAATACTCA	654
Db	219	GlnSerIhisIhisGlnIthrLysProGlnSerSerCysProSerSerAsnGlnAsnThrGln	238
OY	655	GAGATTCOCACCTGCCATTACAGCTGTTCTTCATACAGCAACAAGAACACGATGGTGAAC	714
Db	239	LysSerProProAlaIleAlaLaseSerSerTyrMetProGlyLysIhisSerIleGlnAla	258
OY	715	AGAGATCATAGAGAGGACATGCAGACCCCTGTTCTCCTCAGAGATGCCAGCCTCAGGCACG	774
Db	259	ArgAspIhisIthrGlySerMetGlnIhisCysProAlaGlnArgGlyGlnIleAlaLase	278
OY	775	TGCAGCCCTCAGAAATATATCTGCTCCCTATTAATATACAAAGCTGGAGACCACCTTCCCC	834
Db	279	HisSerPro-----ArgMetLeuProTyrGlnAsnIthrAsnSerGlnLysProAspPro	296
OY	835	AAAAGTCTGATTAAGAAAGATGTCACACACAAATGAATGTACTTGGAGATATACGCCGC	894
Db	297	ThirLysProAspGlnLysAspValTyrPheGlnAsnIleIthrTyrIleGlyLeuTyrSerArg	316
OY	895	CAGGAGTGAAGGAAGCATTCATGAGAGAAACAAAGAGTGTAGTTCTTGGTCCAGAT	954
Db	317	GlnAlaValGlnAspValIleuMetLysGlnAsnIhisAspGlyIthrPheLeuValAlaGAsp	336
OY	955	TGTTTCACAAAATCCAAAGGAAGACCCCTATGTTTGGCTGTGTTTATGAGAACAAAGTC	1011
Db	337	CysSerThirLysSerLysAlaGlnProTyrValLeuValValPheTyrGlnAsnLysVal	356
OY	1015	TACAAATGTAAAAATCCGCTTCTTCGTAGAGAAATACAGATGTTCCCTGGGAGACAGGACGC	1077
Db	357	TyrAsnValLysIleArgPheLeuGlnSerAsnIleInPheAlaLeuGlyIthrGlyLeu	376
OY	1075	AGAGAGATGAGAAAGTTGATTGCATAGTGAAGACATCATCCAAACACTAC	1122
Db	377	ArgGlyAsnGlnMetPheAspSerValGlnAspIleIleGlnIhisTyr	392
RESULT 3			
090Z62			
ID	090Z62	PRELIMINARY:	PRT: 435 AA.
AC	090Z62		(TremBRel. 13, Created)
DT	01-MAY-2000	(TremBRel. 13, Last sequence update)	
DT	01-MAR-2002	(TremBRel. 20, Last annotation update)	

GN CLNK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SOURCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=20029811; PubMed=10562326;
 RA Cao M.-Y., Davidson D., Yu J., Latoro S., Veillelle A.;
 RT "Clunk, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
 stimulated Hemopoietic Cells.";
 RL J. Exp. Med. 190:1527-1534(1999).
 DR EMBL; AF187819; AF14299.1; -;
 DR HSSP; P23727; 2PNB.
 DR MCD; MG1:1351468; Clnk.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PFO0017; SH2; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PSS0001; SH2; 1.
 SQ SEQUENCE 435 AA; 49492 MW; 5CD27EC971F0C0A5 CRC64;

Alignment Scores:
 Pred. No.: 3.24e-95 Length: 435
 Score: 1132.00 Matches: 229
 Percent Similarity: 71.81% Conservative: 41
 Best Local Similarity: 60.90% Mismatches: 102
 Query Match: 56.04% Indels: 4
 DB: 11 Gaps: 3

US-09-856-061-3 (1-1129) x 0902E2 (1-435)

QY	1	TTCCGAGACTTCAGCTCGCCAAAAAGAGGTCATGGCCCTGCATCAATAGTCACAGGC	60
DB	19	PheGlnAsnValSerLeuLeuLysAsnArgSerTyrProSerLeuSerSerAlaLysGly	38
QY	61	CAGTACCCAGAGAGAGTAAACAAAGCCTCTCTTACACTGGGAGAGAAACTTGGCTGCAGTCTG	120
DB	39	ArgCysArgAlaValLeuGlnProLeuProAspHisArgArgAsnLeuAlaGlyValPro	58
QY	121	GATGGAGCAAAAGGCCACAGTGTATGATGACTATGATACCTTGAGCTTCGGATGGAGAGAC	180
DB	59	GlyGlyGlnLysCysAsnSerSerAsnAspArgLysProGluPheGlnLeuLys	78
QY	181	ACATGGCAGCTGATTAATAATTTCCAGCCGCCGCTTAAAGGAATCGAATGAGAGAT	240
DB	79	AlaTyrProSerMetLysIleLeuProAlaArgProIleGlnGlnSerGluTyrAlaAsp	98
QY	241	ACACATATTTCAAGGTTCAGTGAACGACACTCCCTCCGTTAGACACAGGACCTTATC	300
DB	99	ThrArgTyrPheGlnAspThrMetGlnAlaProLeuLeuLeuProProlLysAlaSerVal	118
QY	301	TCCATTGACACAGCCGACCTGTGACACACAG--ACGAGTTGGAGAGTGGACAAACC	357
DB	119	SerThrGlnArgGlnThrArgAspValArgMetThrIleLeuGlnGluValAspLysPro	138
QY	358	ATTTCAGGAGACACGTGAGAAAGCCAAACACTTAAAGAGATGCATCCGTAAGAAAGAACAG	417
DB	139	ThrPheLysAspValArgSerTrnArgPheLysGlyPheLysTyrTrnLysIleAsnLys	158
QY	418	ATTTCCTTTACACCTCTCTCGGCTCTCTTAAACACTTCGGAAGAGTACAAACCTTGGCC	477
DB	159	ThrProLeuProProArgProAlaIleThrLeuProLysTyrGlnProLeuPro	178
QY	478	CCT--GAGCCGAGAGACAGAGCCACCTTATCTCAGAGACACACCTTCCAGAAATC	534
DB	179	ProAlaProProGlnGlnSerSerAlaTyrPheAlaLysLysProTrnPheProGlnVal	198
QY	535	CAGGGAATCCCAATACATTAAGCTTAAGGGGACTTAATGAGGTCCTTGAAGCAGAAAA	594

```

Db 199 GlnArgGlyProArgGlnArgSerAlaLysAspPheSerArgValLeuGlyAlaGlu 218
OY 595 GTTCCTCATACCGAGGAGGAGCCTGATATCAGCATCTCTGTAGAAAACCAAACTACTCAA 654
Db 219 GluSerHisHisIleThrLysProGluSerSerCysProSerSerSerSerLnsnTrGln 238
OY 655 GAGATTCCACTTGCCTTAGCAGCTTCTTCATTCACGACAGCAACGACCAAGCTGTCGAAAC 714
Db 239 LysSerProProAlaIleAlaIleSerSerSerTyrMetProGlyLysHisSerIleGlnAla 258
OY 715 AAGAGTCATAGAGAGGAGGATGACGCCCTTCTCTCCAGAGATGCCAGCTCCAGCCAGC 774
Db 259 ArgAspHisThrGlySerMetGlnHisCysProAlaGlnArgCysGlnAlaIleAlaSer 278
OY 775 TCGAGCCCTCAGCAAAATATACCTGCTTAATATACAGCAAGCTGGAGACCCTTCCGC 834
Db 279 HisSerPro-----ArgMetLeuProTyrGlnValThrHisAsnSerGlyLysProAspPro 296
OY 835 AAAAGTCTGATAGAAAGATGTCCAGCAATGAATGTACATTTGAGAAATACAGCCGC 894
Db 297 ThrLysProAspGlyLysAspValTyrGlnAsnGlnTyrTyrIleGlyLysSerArg 316
OY 895 CAGGAGTGGAGAGGAGCATTCATGAAGAGCAAGAGATGATTTCTGTGTCGAGAT 954
Db 317 GlnAlaValGluAspValLeuMetLysGlnAsnLysAspGlyThrPheLeuValArgAsp 336
OY 955 TGTTCACAAAATCCAGAGAGAGCCCTATGTTTGCTGTGTTTATGAGAACAAAGTC 1014
Db 337 CysSerThrLysSerLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 356
OY 1015 TACATGTAAATCCGCTCTCTGAGAGAGATTCAGCAAGTTCGCTTGGCGGAGCAGACT 1074
Db 357 TyrAsnValLysIleArgPheLeuGlnSerAsnGlnGlnPheAlaLeuGlyThrGlyLeu 376
OY 1075 AGAGAGATGAGAAAGTTGATCTAGTAGAAGCATCATCAGCACTAC 1122
Db 377 ArgGlyAsnGluMetPheAspSerValGluAspIleIleGlnHisTyr 392

RESULT 4
OYGC1 4
ID OYGC1 PRELIMINARY; PRT; 552 AA.
AC OYGC1;
DT 01-MAY-1999 (TREMblrel. 10, Created)
DT 01-MAY-1999 (TREMblrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMblrel. 20, Last annotation update)
DE B cell linker protein BLNK.
OS BLNK OR BASH.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9146381; PubMed=10023776;
RA Ishida M., Kurosaki M., Pappu R., Okawa K., Ronko I., Fu C.,
RA Shibata M., Iwamatsu A., Chan A.C., Kurosaki T.,
RT "BLNK required for coupling Syk to PLC gamma 2 and Rac1-JNK in B
RT cells."
RL Immunol 10:117-125(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99049801; PubMed=9834055;
RA Gotsuka R., Fujimura Y., Mamada H., Umeda A., Morimura T.,
RA Uetsuka K., Doi K., Tsuji S., Kitamura D.,
RT "BASH, a novel signaling molecule preferentially expressed in B cells
RT of the bursa of Fabricius."
RL J. Immunol. 161:5804-5808(1998).
DR EMBL: AF089727; AAD12783.1; -
DR HSSP: P08487; 2PLD.
DR Interpro: IPR002965; P_rich_extensn.
DR Interpro: IPR000980; SH2.

```

```

DR Pfam: PF00017; SH2; 1.
DR PRINTS: PRO1217; PRICHEXTENSN.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS00001; SH2; 1.
SQ SEQUENCE 552 AA; 61822 MW; FB232179BB3BD072 CRC64;

Alignment Scores:
Pred. No.: 5,856-22 Length: 552
Score: 334.00 Matches: 112
Percent Similarity: 37.31% Conservative: 60
Best Local Similarity: 24.30% Mismatches: 157
Query Match: 16.53% Indels: 132
DB: 13 Gaps: 15

US-09-856-061-3 (1-1129) x OYGC1 (1-552)
OY 94 TGGAAAGAACTTGTGCTGAGTCTGATGAGCAAAAGCCACAT----- 141
Db 83 TrpSerAspAspPheAspSerAspTyrGlnAsnProAspGlyHisSerAspSerGluMet 102
OY 142 -----GATGATGACTATGATGACCTGAGCTTCGGATG 174
Db 103 TyrValValProSerGluGluAsnProAspAspSerTyrGlnProProProSerGluGln 122
OY 175 GAGAGCATGGCAGTCATTAATTTTACCAGCCGCTATTAAGCAATCTGAATAT 234
Db 123 GluLysLys-----LysIleProSerSerPheProIleSerArgGlyLysTyr 138
OY 235 GCAGAT-----ACACACTATTTCAGGTTGCAATGACACACCTCCCTCCGCTTAGAC 285
Db 139 AlaAspAsnArgThrSerHisIleGlnLeuProIleAsnLysProLeuProSerThr 158
OY 286 ACCAGAGCTATC-----TCCATTCAGACAGCCGACCTGGAACACA 327
Db 159 ProSerSerAlaLeuProArgProLysLysProSerLeuProSerProAlaAlaLysPro 178
OY 328 CAGACGAGCTG-----GAAAGAGTGGCAAAACCAT 360
Db 179 LysLeuProLeuLysProArgGluCysSerAspAspGluAspAsnTyrIleValProVal 198
OY 361 TCCAGGAGCTCAGAGAGCCAAACATTAAAGAGATGCATCCGTA----- 405
Db 199 AspAsnAspAspAspAsnTyrIleGluProThrGlnSerSerThrProProAlaLys 218
OY 406 -----AGAAAGAACAGATTCCTTACCACTCTCT 435
Db 219 ProProValAlaAsnArgPheMetLysProProAlaLysSerAlaLeuProThrProProLys 238
OY 436 CGGCTCTCATMACACTTCCGAGAGATACCAACCTTGCCCGCTGAGCCGAGAGCAGC 495
Db 239 ProSerLeuAlaSerAspMetGlnGluValTyrGluValProGlnGlnGlnGlnLeu 258
OY 496 AGGCACCTTATATCAGACACACCTTCCAGAGCTCCAGGAATGCCAGT----- 549
Db 259 SerProProProValThrArgPheThrLysPro-----LeuProAlaThrArg 274
OY 550 -----CAGATAAGCTTAAGGAGCTTAAGTAGAGTCTTGAA 585
Db 275 AlaGlnAsnAlaGlnHisSerHisMetHisSerMetThrArgGluSerProLysLeuAsp 294
OY 586 GCAGAAAA-----GTTCTCATACAGAGAGAGCCTGATCAATCAATCATCTG 633
Db 295 AlaSerArgAsnIleLeuProLeuProArgAsnArgLeuHisProLysThrAspHisGlu 314
OY 634 TTAGAAAA-----CAAAATCTCAAGAG----- 657
Db 315 AlaAsnAsnAsnAspGluAsnHisSerPheSerAsnThrGlnGlnSerLysPheProPro 334
OY 658 -----ATTCCACTGGCATTAAGCACTCTCA----- 664
Db 335 GlyAlaAlaProSerProLeuProArgAlaLeuLysLysThrSerAsnAlaValAsnPro 354

```

Oy	685	-----TTCAGGACAGCAGACCACCAAGTGCTGCAAAC	714
Db	355	AlalysProCysLeuProSerAArgspThrPheThrValAsnGluAspLysProThraLa	374
Oy	715	AGAGATCATAGAGAGCGCATGGAC-----CCCTGTTCCTCAGAGATGCCAGCT	765
Db	375	AlAspArgArgArgGlySerSerHisGlnPheProLeuProProLeuProSerGlyThr	394
Oy	766	CCAGCCGAC-----	774
Db	395	ProLysSerSerLeuGlnLysProLeuValLeuProLysValProGluAlaProSerArg	414
Oy	775	-----TGCAGCCCTCACGAAAATAACTACTGCCCTTAATAATACACAAGCTGGAGACCA	825
Db	415	AlaLeuGlyThrSerProHisSerSerIleSerSerIleSerSerThrAla-----	431
Oy	826	CCTTTCCCCCAAAGGTCGTGATTAGAAGAGATGCCACACCAATGAATGGATCTGGACA	885
Db	432	-----AspGlnAspAlaGlyValHisSerLysAlaITryptAlaAlaIthr	446
Oy	886	TACAGCCCGCAGCGCATGTGAAGAGCATTCATGAAGAGACAAGAGATGTAGTTCTTG	945
Db	447	CysAspArgLysThrAlaGlnAspAlaLeuTyArgSerAsnLysAspLysSerPheLeu	466
Oy	946	GTCGCAATATTGTTCCACAAATCCACAAGAGAACCCCTATGTTTGCGTGTGTTATAG	1005
Db	467	IleArgLysSerSerGlyGlnAspSerArgGlnProTyrThrLeuValAlaPheTyrAsn	486
Oy	1006	AACAAGTCTACATGTAAAAATCCGCTCTCGTAGAGAGATCACACAGTTTGCCCTGGGG	1065
Db	487	ArgArgValLysTrsnIleProIleArgPheIleGluSerThrArgGlnThrAlaLeuGly	506
Oy	1066	ACAGGACTCAGAGAGATGTGAAGATTGATTCTAGTGAAGACATCATCGACACTACAG	1125
Db	507	ArgGlnLysCysGlyGlnArgPheAspSerValAlaGluIleValGlnAsnHisGln	526
Oy	1126	AAAT 1128	
Db	527	HIS 527	

RESULT 5
Q9DGC07 PRELIMINARY; PRT: 530 AA.

AC	09DGC07:	
DT	01-MAR-2001 (TREMBLE).	16, Created)
DT	01-MAR-2001 (TREMBLE).	16, Last sequence update)
DE	01-MAR-2002 (TREMBLE).	20, Last annotation update)
GN	SLP76.	
OS	Gallus gallus (Chicken).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;	
OC	Gallus.	
OX	NCBI_TaxID=9031;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=THYMUS;	
RX	MEDLINE=20451093;. PubMed=1093915;	
RA	Ishai M., Kurosaki M., Inabe K., Chan A.C., Sugamura K., Kurosaki T.,	
RT	"Involvement of LAT, Gads, and Grb2 in Compartmentation of SLP-76 to	
RT	the Plasma Membrane."	
RL	J. Exp. Med. 192:847-856(2000).	
DR	EMBL, AF226988: AAG18493.1; -.	
DR	HSPF; P12931.1SHD.	
DR	InterPro; IPR001660; SAM.	
DR	InterPro; IPR000980; SH2.	
DR	Pfam; PF00017; SH2_1.	
DR	PRINTS; PR00401; SH2DOMAIN.	
DR	Prodom; PD000093; SH2_1.	
DR	SMART; SM00454; SAM_1.	
DR	SMART; SM00252; SH2_1.	
DR	PROSITE; PS50001; SH2_1.	

Seq	SEQUENCE	530 AA	59580 MM	CF6DBCBE8D97909A	CRC64:
50	ALIGNMENT SCORES:				
	Pred. No.:	1,53e-20	Length:	530	
	Score:	318.50	Matches:	107	
	Percent Similarity:	37.59%	Conservative:	55	
	Best Local Similarity:	24.83%	Mismatches:	143	
	Query Match:	15.77%	Indels:	126	
	DB:	13	Gaps:	15	
US-09-856-061-3 (1-1129) x Q9D607 (1-530)					
QY	120	GGATGGAGCAAAAGCCACATGATGACTATGACTATGATGACCTTGAGCTTGGATGTGAGAGA	179		
DB	103	GLYTPSERSERPHEASPLP--ASPASPSPYTYGLUSERPROASPASPAPSPGLNLIUY	122		
QY	180	GACATGGCAGTCGATTTAAATTTTACACAGCCCGCCCTATAAGAAATCTGATATGTCAGAGA	239		
DB	122	SGLUSPGLTUNLAPSPYTYGLUSERPROTHGLUGLTFHGLINGLUALAGLUNHISPSPE	142		
QY	240	TACACACATTTTCAAGGTTCGATGACACTCCCTTCCTCCCTAGACACAGGACCTCTAT	299		
DB	142	RASPLLYTYR-----GLUPROPROPSERASNAHNSPGLUNLHISHI	157		
QY	300	CTCCATTGGACAGCCGACCTTGGAGACACACAGACAGGTGTGAAAGATGACACAACCCAT	359		
DB	157	SASNAVAILLEPHEPROLALYSERLEUALLASNAHNTHTASPYRTILEASPARPRO--	176		
QY	360	TTCCAGGGACGTCAAGAGCCAAAACATTAAGAAGATGATCCGTAAAGAAAGAAAGAT	419		
DB	177	-----ProTHSerArgSerHisGlnPr	185		
QY	420	TCCTTTACACCTCTCTCGGCT	443		
DB	185	OPROVALPROPROGINARPGPROGLYPROSERPROLAPROLASERHEGLYARGL	205		
QY	444	CATAACACTTCCGAAGAAGTACCAACCTTG--	474		
DB	205	YALASERLEUPRO--ALAPHEPROPROLEUPLROGLYASNAHNSPARGHANVALYSPR	224		
QY	475	-----CCCCCTGAGCCGGAGAGACAGG--	498		
DB	224	OSERLYSPROPROLAPROSERLLEASPARGSETHLYSPROLEUASPARGLLEUGL	244		
QY	499	CCACCTTTATCTCAGACACACACCTTTCCA---GAAGTCCAGGGAATGCCAGTCCAGAT	554		
DB	244	YPROPHROHEGLUARGLUASNPPOVALPROGLYARGLYSPROGLYHISPROGLIUNLYSLE	264		
QY	555	A-----AGCTTAAGGACCTTAAGTAGAGGTCCTTGAGACGAAANAAGTTCCTCTAA	605		
DB	264	ULEUTHRPROGLINLEUARGALALEUGLYGLNGLUE-----ALAWETMETPROLYSPR	282		
QY	606	CCAGAGGAAGCCTGGAATCAACTCACTCTGTGAAAAACCAAAAT-----ACTCAAGAGAT	659		
DB	282	OPROVALPROPROSERASPARTYTYGLUARGLYASNPBROSERPROLEUARGLYSGNLI	302		
QY	660	TCACCTTGCCATTACAGACTTTTCATTTCACAGACAAGCAACACAGTAGTGCAAAAACAGAGA	719		
DB	302	EPROVAL-----LYSGINGLYTPRALGLINGLUNLYSARGPROGLIUNGLIUNLWS	319		
QY	720	TCATAGAGAGGAGCATGACACCCCTGTTCTTCCTCAGAGATCCAGCCTCCAGCAGCTCAG	779		
DB	319	PHIS-----ILEPROGLNARGLALVALPROGLINLISERLEUPR	332		
QY	780	CCCTCAC--GAAATATACTGCCCTATAAA--	807		
DB	332	OPROTYRSESRASNTHTPHEPROSERLYSERLLELYSALAPROPOLYSPROGLYSE			

```

Db      372  oarGpheTroLeuGlYasnAsnSerArgSerProSerArgGlYThrAlaAspLeuArgtr  (|||||
QY      825  ACCTTTCCCC-----AAAAGCTCTATGAAAGA 854  ||||| |||
Db      392  opTroLeuprTrolerProSerArgGlYThrAlaAsnGlYThrAsnThrcuLnuLAspGluAs  (|||||
QY      855  TGTCACACCAATGAATGATGACATTCGACAAATACAGCCGCCAGACAGTGAACAGGCATT 914  ||||| |||
Db      412  pSerLeuAsnAspLnuTrpTryValAlaIaIaTryLieserArgProGluAlaIaIaIaIe  (|||||
QY      915  CATGACGAGAACAAAGATGGTAGTTCCTTGTCGCGAGATTGTGCCAACAAATCCAAAGA 974  ||||| |||||
Db      432  uATrGluAlaIeAsnLysAspGluYThrPheLeuValAlaArgAspSerArgYThrThrTh  (|||||
QY      975  AGAGCCCTATGTTTGCTGCTGTGTTTGTGAGAACAAAGTGTACAAATGTAATAATCCGGTT 103  ||||| |||||
Db      452  rHAsProGlyValLeuMetValLeuYrYrLysAspLysValAlaYrAsnIleGlYnLeuArgY 472  ||||| |||||
QY      1035  CTTGGACAGCAAAATCCAGCGATTGCCCTGGGAGCAGAGCATCAGAGCATGACAAAGTTGA 109  ||||| |||||
Db      472  rGlnGlYnGlYAspGlnThrYrYrLeuLeuGlYnTrGluYrLeuYrGcLYrGcLnuAspPhe 492  ||||| |||||
QY      1095  TTTCAGTGAAGACATGATCTGCAACAACATCAAG 1125  ||||| |||||
Db      492  rSerValAlaAlaAsnIleLeuLeuAspPrrYrPheGln 502  ||||| |||||

```

RESULT 6	092010	PRELIMINARY;	PRT;	525 AA.
ID	092010			
AC	092010;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TREMBLrel. 20, Last annotation update)			
DE	SLP-76 adaptor protein.			
CN	SLP76.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus			
ON	[NCBI_TaxId=10116;			
RX	[1]			
RP	SEQUENCE FROM N.A.			
RA	Fujii Y., Goltshuka G., Wakahara S., Hara T., Nakao T.;			
RT	"Molecular cloning of SLP-76, a 76-kDa adaptor protein in RBL-2H3."			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBD databases.			
DR	EMBL; AB072980; BAB1779.1; -			
DR	InterPro: IPR000980; SH2.			
DR	Pfam: PF00017; SH2; 1.			
DR	ProDom: PD000093; SH2; 1.			
DR	PROSITE: PS00001; SH2; 1.			
SEQUENCE	525 AA; 59521 MW; 0B1CDD40DABE7372 CRC64;			

Alignment Scores:	
Pred. No.:	2,34e-20
Score:	316.50
Percent Similarity:	36.81%
Percent Local Similarity:	25.94%
Query Match:	15,67%
DB:	11
Length:	525
Matches:	117
Conservative:	49
Mismatches:	112
Indels:	173
Gaps:	19

US-09-856-061-3 (1-1129) x Q920L0 (1-525)

[illegible]

Qy	238	GATTCACACATATTTCAAGGTTGCAATGGACACTCCCTCCCTTCCTTGACACACGACCTCT	297
Db	166	-----	166
Qy	298	ATCTCCATTGGACAGCCGACCTGGACACACAGACAGAGTTTGGAAAGCTGACAAACCC	357
.Db	167	-----Proasnthrhansermet-----TyllasparigPro	177
Qy	358	ATTTCAGGAGCTGCAGAGGCCAAACATTAAAGACATGCATCCGTAAGAAACACAG	417
Db	178	ProthcglyuylvalseringlnPro-----	186
Qy	418	ATTCTTTACCACTCTCCGGCTCTCATACACTTCCGAAAGATCAACACCTTGGCC	477
Db	187	---ProvalProproglnarPrometalaaleuPro-----ProleuPro	201
Qy	477	-----	477
Db	202	ThhGlyArghsnhiserProasnhisglugluProserArghsnArghsnhisllyTh	221
Qy	478	-----CCTGAGCCG-----GAGAGCAGACAGCCACTTATGTCAG-----	513
Db	222	AlalyuylProalarProserlileasparigserThrllyProProleuAsparigSerleu	241
Qy	513	-----	513
Db	242	AlArProleuAsparigluProleuengllyllyllylProserArlylProserAla	261
Qy	514	-----AGACACACCTTTCCAGACACTCCAG-----GGAATGCCCATGCAGATACAG	558
Db	262	ProleuglyArghlnhlsleuProlyslleolnlylProProleuProProalameAr	281
Qy	559	TTAAGGCACTTAAGTCAG-----GTCCCTGACACAGAAAAA-----GTTCCTCAT	603
Db	282	ArghnlsghlArghsnlArghlleolylProleuThrlAlArghlylProProvalProArgh	301
Qy	604	AACCCAGGAGAACCCGAA-----	621
Db	302	hlsghlyArghlyProAsparigArghArghsnAspSluAsparvalhlsghlnarProleu	321
Qy	622	-----TCAACTCATGTGTTAAACAAACCAAAATCTCA-----	654
Db	322	ProhlsProserleuProSerMetSerSerhlnhrhProserArghserThrllylPro	341
Qy	655	-----GACATTCACCTGCCATTAGCAGTTCTTCATTC	687
Db	342	SerProlyAsnthrhProleuProhlsleuProglAlaleSerghlsnAlle	361
Qy	688	ACGACAAACACACACAGTGTGCAAAACAGATCATAGAGAGGCGCATGGACGCCCTGTCT	747
Db	362	-----GlyPhnglnInserAlaser	368
Qy	748	-----CCTCAGACATGCCACCTCCAGCCAGTCCAGCCCTCAGCAAAAT-----	792
Db	369	leuProserTyrlpheSerghlnghlyProserSerArghProvalArghsnghluvalArgh	388
Qy	793	ATATCGCCCTTAAATACACAGAGCTGAGACACACTTCCCAAAAGCTCATATGAAG	852
Db	389	AsnleuProleuProvalProAsn-----ArghProglnProProserProglInghluInlu	407
Qy	853	GATGTCACAGCACAATGAATGTGATCAATGGAGATACAGCCGCCAGCATGGAGAACGCA	912
Db	408	SerProleuAspghlnghlylProghlylserTyrlleThArghProghlnghlnghlnAla	427
Qy	913	TTTCATGAAGGAGAACAGATGTGATGTTTGTGTGCGGATGTTTCCACAAATCCAAAG	972
Db	428	leuArghlyllesnghlnAspghlylThrhlelevalArghsSerSerlylshlylval	447
Qy	973	GAAAGACCTATGTTTGGCTGTGTTTATGAGAACAAAGTCTACATGTTAAATCCGC	1032
Db	448	AsnAsnProghlylAlleuMetValleuTyrllylAspghlylvalTyrlAsnlnleghlnleArgh	467

```

OY 1033 TTCCTGAGAGCAATCAGCTTTCCTGGGACAGACGATCAGAGATGAGATT 1092
DB 468 TyrgIngluGlusergInValTyrluLeuGlTyrluLeuArgLyLysGluAspHe 487
OY 1093 GATTCAGTAAGACATCATCGACACTACAAG 1125
DB 488 LeuSerValSerAspIleLeuSerPyrPheArg 498

RESULT 7
O922M0 PRELIMINARY; PRT; 533 AA.
AC O922M0:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Lymphocyte cytosolic protein 2.
GN LCP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; BC006948; AH06948.1;
DR MGD; MGI:1321402; LCP2.
DR InterPro; IPR000980; SH2.
DR Pfam; PF00017; SH2; 1.
DR ProDom; PD000093; SH2; 1.
DR PROSITE; PS00001; SH2; 1.
SQ SEQUENCE 533 AA; 60238 MW; 50AEAD25EF0AD01 CRC64;

Alignment Scores:
Pred. No.: 3.58e-20 Length: 533
Score: 314.50 Matches: 121
Percent Similarity: 40.94% Conservative: 44
Best Local Similarity: 30.02% Mismatches: 169
Query Match: 15.57% Indels: 69
DB: 11 Gaps: 18

US-09-856-061-3 (1-1129) x O922M0 (1-533)
OY 58 GGCCAGTACAGAGATGAGACAAGCCTCTAGACTGGGAAAGAACTTTCGTCGATC 117
DB 126 GlyAspTyrGluSerProAsn-----GluGluGluGlnAlaLeuVal 139
OY 118 CTGGATGGAGCAAAAGGCCACAGTGATGACTATGATGAGCCCTCGATGAA 177
DB 140 AspAspAlaAla-----AspTyrGluProProProSerAsnAsnGlu 153
OY 178 GAGACATGGCAGTCGATTAATAATTTACAGCCCGGCGCTATA-----AAGAAATCGAA 231
DB 154 GluAlaLeuGlnSer--SerIleLeuProProAsnSerPheHisAsnThrAsnSerMet 172
OY 232 TATGCAGAT--ACACACTATTTCAGGTGCAATGAGACACTCCCTT--CCGTTAGAC 285
DB 173 TyrIleAspArgProProThrGlyLysValSerGlnGlnProProValProProLeuArg 192
OY 286 ACCAGGACCTCTATCTCCATTGGACAGCCGACCTGGAAACACACAGAGCTTGAAGA 345
DB 193 ProLysProAlaLeuPro--ProLeuProThrGlyArgAsnHisSerProLeuSerPro 211
OY 346 GTGAGACAACCCATTTCCAGGAGCGTCAGAGCAAAACCTTAAGAGATGATCCGTA 405
DB 212 ProHisProAsnHisGluGluProSerArgSerGlyAsnAsnLys----- 226
OY 406 AGAAAGACAAGATTCCTTTACCA-----CCTCCGCGGCTCTC 444
DB 227 ---ThrlAlaLysLeuProAlaProSerIleAspArgSerThrLysProProLeuAspArg 245
OY 445 ATAACTTCCGAAGAAGTATCCAAACCTTG-----CCCCCTGAGCCGGAG 489

```

```

DB 246 SerIleAlaProLeuAspArgGluProPheIleLeuGlyLysProProPheSerAsp 265
OY 490 AGCAGCAGGCCACCTTTATCTCAGAGACACACCTTTCCAGAGATCCAG-----GGATG 543
DB 266 LysProSerAlaProLeuGlyArgGlnHis--LeuProLysIleGlnLysProProLeu 284
OY 544 CCCAGTCAGATAGCTTAAGGACTTAAGTGAGCTT-----GAGCA 588
DB 285 ProProAlaMetAspArgHisGluArgAsnGluArgLeuGlyProValThrThrArgLys 304
OY 589 GAAAAGTTCCCTCATACACAGAGAAAGCCGTGATCACTCATCTGTAGAA----- 639
DB 305 ProProValProAlaGlnHisGlyArgGlyProAspArgArgGluAsnAspGluAspArgVal 324
OY 640 AACCAAAATACTCAAGACATTCGCACTTGCATTACAGTTCTTCAATCAGCAAGCAAC 699
DB 325 HisGlnArgProLeuProGlnProSerLeuProSerMetSerSerAsnThrPheProSer 344
OY 700 CACAGTGTCA----- 711
DB 345 ArgSerValGlnProSerSerLysAsnThrPheProLeuAlaHisMetProGlyAlaPhe 364
OY 712 AACAGAGATCATAGAGAGGAGCATGCAAGCCCTGTCT-----CTCAGAGATGCCAGCT 765
DB 365 SerGluSerAsnIleGlyPheGlnGlnSerAlaSerLeuProProTyrPheSerGlnGly 384
OY 766 CCAGCCAGCTGCAGCCCT-----CAGCAAAATATCTGCCCTATTAATACACAGC 816
DB 385 ProGlyAsnArgProProLeuArgSerGluGlyArgAsnLeuProLeuProValProAsn 404
OY 817 TGGAGACACCACTTTCACCAAGGTCGTGTAGAAAGATGTCCAGCAATGATGTTAC 876
DB 405 ---ArgProGlnProProSerProGlyGluGluGluThrProLeuAspGluGluTyr 423
OY 877 ATTGGAATPACAGCCCGCAGGACGTGAAAGGCACTTATTAAGAGAAACAAGATGTT 936
DB 424 ValSerTyrIleThrArgProGluAlaGlnAlaLeuArgLysIleAsnGlnAspGly 443
OY 937 AGTTCTTGTCGAGATGTTCACCAAAATCCAAAGAGAGAGCCCTATTTTGGCTGTG 996
DB 444 ThrPheLeuValArgAspSerSerLysIleThrAlaAsnAsnProTyrValLeuMetVal 463
OY 997 TTTTATGAGACAAGCTACATGTAATAATCCGCTTCGAGAGAGATCAGCATTT 1056
DB 464 LeuTyrLysAspLysValTyrAsnIleGlnIleArgTyrGlnGluGlnSerGlnValTyr 483
OY 1057 GCCCTGGGACAGACTCAGAGAGATGAGAGATTTGATTGACTACAAACATCATCGAA 1116
DB 484 LeuLeuGlyThrGlyLeuArgGlyLysGluAspPheLeuSerValSerAspIleLeuAsp 503
OY 1117 CACTACAAG 1125
DB 504 TyrPheArg 506

RESULT 8
O75499 PRELIMINARY; PRT; 433 AA.
ID O75499:
AC O75499:
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE B cell linker protein BLNK-S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=98001722; PubMed=9341187;
RA Fu C.; Chan A.C.;
RT Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
RT receptor activation.

```

QY	60	CSAATACCAAGCATGAACAAGCCSTCTTTCAGACTGGGAAAGAATT-----	108
Dd	57	ProLaAsrpluGluglInglntPrSsrAsp-AspPhaAsrSerIAspryltIdasnPrfAs	76
QY	109	-----CSTGCAGTCTCGATGTGGAGCAAAAAGGCCACAGTAGATGACTA	152
Dd	76	rGIunHisSerAsrSergIumEtlytValImeStrpDlaIgLIuIsnLlaAsrpsrEty	96
QY	153	TGATGACCSTGAGCTGGSGATGGAGAGACATGGSCAGTGCATTAATAATTTCACAGCCC	212
Dd	. 96	rGIuPrPrtProvalIdluIgltu-----ThlrArgProvaIlnIsrProIale	112
QY	213	GCCATTAAGAAGAACTCTGAATATGCAGATCACACATACTATTCAAGTTGCAAATGGACATCC	272
Dd	112	wPrOrhAlaIatgSlLyutIurIttleAsrpaNalq-----SererGIarhI	128
QY	273	CSTTCCSTTAAGACACAGGACCTCTATCTTCATTTGACAGACCCGACCTGAAACACAGAC	332
Dd	128	sSerPrOrhPrtPrtPrtSerIythrLeuProser-----LysProSerTrp-----	142
QY	333	GAGGTTTGAAAGACGTGCAACAACCCATTTCTCAGGAGCGTGAAGAACCCAAACATTAAAGS	392
Dd	143	-----ProserGIuLyAlatIatgIueThIserThrleuProal	155

OY	393	AGATGCAATCCGAAAGAAAGCAACAGATTCTTTACGACGCTCCGCGCTCTATA-----	447
Db	155	aleuththalaleuqnllysrprogluvalpro--rplysrprolyseleuqnllysls	174
OY	448	-----ACACTTCCGAAGA	461
Db	174	pcgluallasptryvalvalprovalgluaspasnasrgluasntryllehtisprrothrgl	194
OY	462	GTACCAACCCTTGGCCCTTGAGCCG-----GAGAGCAGCAGGCC	500
Db	194	userseerseerproproprogluylsrglyargasnserylalatprrllythrlysserpr	214
OY	501	ACCTTTATCTAGAGACACACCTTTCCAGAAAGTCACAGAGATGCCCGCATAG-----	552
Db	214	oprorolalalaproserproleuprorolalaglylylsysrprothrhrproleuyl	234
OY	553	-----ATAGCTTAAAGGACTTAAGTAGAGCTCTTGAGAGCAAAAAGTTC	599
Db	234	sthrthrprovalalaserenginginsanallasersevalyrglyugluysprollepr	254
OY	600	TCATATCACAGAGAAAGCTGTAATAACTCATCTGTGTGAAAACCAAAATCTCAAGAGT	659
Db	254	o---Alaqluylrghlsarglylserghlnlsarglylnl-----	266
OY	660	TCCACTTCCCATTTAGACGCTTTGATTCACAGACAAAGCAACACAGATGTGCAAAACAGAGA	719
Db	267	-----Alavalglinsrprovalphe-----	273
OY	720	TCATAGAGAGGAGCATGCAGCCCTGTCTCTCAGAGATGCCAGCTCCAGCCAGCTCAG	779
Db	274	-----proproalalnllysgl	279
OY	780	CCCTCACGAAATATAGTACGCCCTAT--AAATACACA--AGCTGGAGACA-----	825
Db	279	nllehtislnlysrproleuprorolalgrhethlglylglylsanprothrlyals	299
OY	826	-----CCTTTCCCAAAAAGTGTGATAGAAAGATGTCCACACATGAA-----	870
Db	299	pcglproleuproserpheserseasnserrthlleseerlglnlvalaglyalle	319
OY	871	-----TGGACACTTGGAGAAATACAGCCGACAGGACGTGAGAGAGCATTCATGAA	920
Db	319	uclysrprothrprlyrlyalaglyalacysasrparlgllysserlalaglylnlaleuhtsrr	339
OY	921	GGAGACAAGAGATGATGTTCTTGTGTCAGAGATTGTTCCACAAATCCAGAGAGAGCC	980
Db	339	gserrlnlysrpdllyserphelneuillalarglysserseerghlnlsasrserlysglnpr	359
OY	981	CTATGTTTGGCTGTGTTTATATGAGCAAAAGTCTCAATGTAATAATCCGCTCTCTGGA	1041
Db	359	otylthrlleuvalvalpheerhenslnlysaargvalylasnllyserprovalrthelheol	379
OY	1041	GAGAAATCAGAGATTTGCCCTGGGACAGAGACTCAGAGAGATGAGAGATTGATTCAGT	1100
Db	379	ualathrlysglnllyrlyalaleuqlylarygllylsasnglylgluglyltyrphedylserva	399
OY	1101	AGAAAGCATCTATCGAACATCAAGAAAT	1128
Db	399	lalaglyllellylalarghasnhtisglnhts	408
RESULT 9			
ID	075498	PRELIMINARY:	PRT: 456 AA.
AC	075498:		
DT	01-NOV-1998	(TEMBLrel. 08, Created)	
DT	01-NOV-1998	(TEMBLrel. 08, Last sequence update)	
DT	01-MAR-2002	(TEMBLrel. 20, Last annotation update)	
DE	B cell linker protein BLNK.		
OS	Homo sapiens (Human).		
OC	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:		
OC	Mammalia: Euthera: Primates: Catarrhini: Homiidae: Homo.		
OX	NCBI_TaxID=9606:		
FN	[1]		

ID	Q8WV28	PRELIMINARY;	PRT;	456 AA.
AC	Q8WV28:			
DT	01-MAR-2002 (Tremblrel. 20, Created)			
DT	01-MAR-2002 (Tremblrel. 20, last sequence update)			
DT	01-JUN-2002 (Tremblrel. 21, last annotation update)			
DE	B-cell linker.			
OS	Homo sapiens (Human).			
CC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi			
CC	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=TONSIL;			
RA	Strausberg R.;			
RL	Submitted (DEC-2001) to the EMBL/Genbank/DBJ databases.			
DR	EMBL; BC018906.1; ANH18906.1; ..			
DR	InterPro; IPR000950; SH2.			
DR	Pfam; PF00017; SH2; 1.			
DR	ProDom; PD000093; SH2; 1.			
DR	SMART; SM00252; SH2; 1.			
DR	PROSITE; PS50001; SH2; 1.			
SQ	SEQUENCE 456 AA; 50465 MW; ED6D424A035D1792 CRC64;			

Alignment Scores:	
Pred. No.:	456
Length:	7.56e-18
Score:	289.00
Percent Similarity:	38.58
Best Local Similarity:	25.63
Query Match:	14.31
DB:	4
	Indels:
	Gaps:
	17

OY	127	GCAAAAGGCCACACTGATGTAGTATGATGACCTTGACCTTCGGATGGAAAGACATGG	168
Dd	88	AleluclunelnalnalaabphbserTyrglunProproProlglunglnu-----	1040
OY	187	CAGTCGATTAAATTTTACCAGCCCGCCTATAAAGAATCTGAATTCGAGATACACAC	246
Dd	105	---ThrArgProValHISProAlaLeuProPhenAlargIlygIutyry-----	119
OY	247	TATTTCAGGTTCGATGACACTCCCTCCCTTCGGTAGCACACAGGACCTGTATCTCCATT	306
Dd	120	-----IleaspasnArgSerSer-----	125
OY	307	GCAGACCCGACCTGGACACACAGACGAGGTGGAGAACCAACCATTTCCAGG	366
Dd	126	-----GlnArgHisSerProProPheSerIlys-----	134
OY	367	GAGCTCGAAGCCCAAACATTAAAGAGATGATCCGTAAAGAAAAGACAAGATTCCTTA	426
Dd	135	ThreLeuproSerIys-----ProSerTrpProSerGlunylalargIeu	149
OY	427	CCACCTCCTGGGCTCTCATACACTTCCGAAAGAAGTACCAACCTTGCCCCCTGACCG	486
Dd	150	ThrSerThrLeuProAlaLeuThrAlaLeuGlnLysProGln--ValProPolysPro	168
OY	486	-----	486
Dd	169	LysGlyLeuLeuGlnuAspGluAlaAspTyrValProValGlnuAspnAspGluAsn	188
OY	487	-----GAGACGACGACGCCA-----CCTTATCTCAGAGA	516
Dd	189	TyrlleHisProthrGlnuSerSerSerProProGlnuLysAlaPrometValasnaArg	208
OY	517	CACACCTTCCACAAGTCGCG-----GGAATGCCCAGTCAGATTAAGCGAC	567
Dd	209	SerThrLysProAsnSerSerThrProAlaSerProProGlnuThralaSerGlyarGaen	228
OY	568	TTAAGTAGAGTCTCTGACAGAGAAAAA-----GTTCTCAT	603
Dd	229	---SerGlyAlaetrpIunThrLysserProProProAlaAlaProserProLeuProArg	247

QY	604	AACCCAGAGAGACCCCTT-----GAATCAACATCATGTTGTACAAAACCAAAATACT	651
Db	248	AlaGlyLysLysProThrThrProLeuLysThrThrProValAlaSerGlnGlnSer---	266
QY	652	CAGAGATTCACATTCCTCCATTAGCAGTTCCTTATTCACGACAGCAACCAACAGTGTCAA	711
Db	267	-----AlaSerValLysGlnLysProIlePro	277
QY	712	AACAGAGATTCATPAGAGAGGC-----ATCCAGCCCTGTTCTCCTACAGATGCCAGCT	765
Db	278	AlaGlnArgHisArgGlySerSerHisArgGlnLysValGlnInsProValIlePro	297
QY	766	CCAGCCACCTCCAGCCCTCAGCAAAATATGTCGCCCTAT---AAATACACA---AGCTGG	819
Db	298	ProAlaGlnLysGlnIleHisGlnLysProIleProLeuProIleArgPheThrGlnGlyGly	317
QY	820	AGACCA-----CCTTCCCAAAAAGTCTGATGAAAGAGATCCACAGCAAT	867
Db	318	AsnProThrValAspGlyProLeuProSerPheSerSerAsnSerThrIleSerGln	337
QY	868	GAA-----TGTCATATTGGAGAAATACGCCCGCAGGACAGTGGAA	906
Db	338	GlnAlaGlyValLeuCysLysProIlePrtYrAlaGlyAlaCysAspArgLysSerAlaGln	357
QY	907	GAGCATTCATGAAGAAGAACAGATGTAGTTCTTGTCGCGACATGTTTCCACAAA	966
Db	358	GlnAlaLeuHisArgSerAsnLysAspGlySerPheLeuIleArgLysSerSerGlnHis	377
QY	967	TCCAAAGGAGAGCCCTTGTTGGCTGGCTGTTTATGAGAACAAAGTCTACAAATGTAAA	1028
Db	378	AspSerLysGlnProTyrThrLeuValAlaPheHisAsnLysArgValYrAsnIlePro	397
QY	1027	ATCCGCTTCCTGTGAGAGGAATCAGCAGTTCCTCCCTGGGACAGACTCAGAGAGATGAG	1088
Db	398	ValArgPheIleGlnAlaThrLysGlnTyrAlaLeuGlyArgLysLysAsnGlyGln	417
QY	1087	AAGTTGATTCAGTGAAGACATCATCGAAGCTACAGAAAT	1128
Db	418	TyrPheGlySerValAlaGlnIleIleArgAsnHisGlnHis	431
RESULT 11			
Q9OUN3	ID	PRELIMINARY;	PRT: 457 AA.
O9OUN3	Q9OUN3		
DT	01-MAY-2000 (Tremblé, 13, Created)		
DT	01-MAY-2000 (Tremblé, 13, Last sequence update)		
DT	01-MAY-2002 (Tremblé, 20, Last annotation update)		
DE	SH2-containing leukocyte protein 65 (Lymphocyte antigen 57).		
GN	SUP-65 OR BASH OR LY57.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BALB/C; TISSUE-LYMPHOID;		
RX	MEDLINE=9837271; PubMed=9705962;		
RA	Wienands J., Schwelkert J., Wollschied B., Juma H., Nielsen P.J.,		
RA	Reich M.;		
RT	"SUP-65: A new signalling component in B lymphocytes which requires		
RT	expression of the antigen receptor for phosphorylation."		
RL	J. Exp. Med. 188:791-795(1998).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-BALB/C; TISSUE-LYMPHOID;		
RA	Wienands J., Larbolette O., Reich M.;		
RT	"Evidence for a preformed transducer complex organized by the B cell		
RT	antigen receptor."		
RL	Proc. Natl. Acad. Sci. U.S.A. 938:7865-7870(1996).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RA	Okamoto T., Hayashi K., Tsuji S., Gotsuka R., Kitamura D.;		
RT	"BASH: B lymphocyte adaptor protein containing SH2 domain."		


```

Oy      820  ---GAGCAGCACTTCCCC-----AAAAGCTCGATGAAAGAAGATTC 858
           :::|:|:|:|:|  |||
Db      309  ProLysProProLysProThrThrValAlaAsnAsnSerMetAsnAsnMetSerLeu 328
           ::::|:|:|:|:|  |||
Oy      859  CAGCACATGAATAGTCATTGGATGAGAAATACAGCCCGCAGCAGTGAAGAGCATTCATG 918
           |||  |:|:|:|:|  ||:|:|  |||  |||  ::
Db      329  GlnAspAlaGluTyrPylTrpGlyAspIleSerArgGluGluValAsnGluLys---Leu 347
           GlnAspAlaGluTyrPylTrpGlyAspIleSerArgGluGluValAsnGluLys---Leu 347

Oy      919  AAGGAGAAACAAGATGGTACTGTTCTTGGTCCGAGATTTGTTCCACAAAATCCAAGGAAAG 978
           :::::  |:|:|:|:|  ||:|:|  ||:|:|  ||:|:|  ::
Db      348  ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp 367
           :::::  |:|:|:|:|  ||:|:|  ||:|:|  ||:|:|  ::

Oy      979  CCTATGTTTGGCTGTGTTTATAGAACAAAGTCTACAAATGTATAA---ATCCGCTTC 1035S
           |||  |||  |||  |||  |||  |||:|:|
Db      368  ---TyrThrLeuThr-----ProArgLysGlyGlyAsnAsnLysLeuIleLysIle 383
           |||  |||  |||  |||  |||  |||:|:|

Oy      1036  CTGAGAGGAATACAGCAGTTGCCCTGGGAGCAGAGCTCAGAGAGATACAGATTGAT 1095S
           ||:|:|:|  :::::  ||:|:|  |||  |||:|:|
Db      384  PheHisArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheAsn 399
           ||:|:|:|  :::::  ||:|:|  |||  |||:|:|

Oy      1096  TCAGTAGAAGACATCATCAGAACTACAGAAAT 1128
           |||:|:|  :::::|:|  |||:|:|:|:|:|
Db      400  SerValValGluLeuIleAsnHisTyrArgAsn 410
           SerValValGluLeuIleAsnHisTyrArgAsn 410

```


GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:14:41 ; Search time 77.294 Seconds
(without alignments)

5933.817 Million cell updates/sec

Title: US-09-856-061-1

Sequence: 1 acgagagcccaactgccagcag.....taaaaaaaaaaaaaaaaa 1721

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 908470 segs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=framer-n2p.model -DEV=alp
-Q=/cgn2_1/USPTO.spool/US09856061/rnuc_21042003_113018_3728/app_query.fasta.1.3150
-DB=A.geneseq.101002 -QMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosun62 -TRANS=human.cd1
-LIST=45 -DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MTLEN=0 -MAXLEN=2000000000
-USER=US09856061.ecgn_1.1_209_6runat.21042003_113018_3728 -NCPU=6 -ICPU=3
-NO_XLPAY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq.101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2316	75.8	435	22	AA881070	Murine mast cell-s
2	1370.5	44.8	428	23	AAE22608	Human MIST protein
3	1316.5	43.1	428	23	AAE22609	Human MIST splice
4	1132	37.0	376	22	AA881071	Human mast cell-sp
5	1109.5	36.3	353	23	AAE22610	Human MIST splice
6	326	10.7	456	22	AA893406	Human protein sequ
7	319	10.4	449	19	AAW59866	Amino acid sequenc
8	319	10.4	456	19	AAW59865	Amino acid sequenc
9	287.5	9.4	457	19	AAW59867	Human SLP-76 SH2 d
10	260	8.5	95	20	AAW70587	Human secreted pro
11	213	7.0	84	21	AA603152	Peptide #5653 enco
12	212	6.9	46	22	AB838147	Peptide #5353 enco
13	212	6.9	46	22	AB823354	Human brain expres
14	212	6.9	46	22	AAW58783	Human bone marrow
15	212	6.9	46	22	AAW71296	Peptide #5407 enco
16	212	6.9	46	22	AAW18973	Peptide #5612 enco
17	212	6.9	46	22	AAW31575	Human peptide enco
18	212	6.9	46	23	ABG41096	Novel human diago
19	195.5	6.4	503	22	ABG22854	Novel human diago
20	165.5	5.4	474	18	AAW15253	Human brain-specif
21	165.5	5.4	594	18	AAW15256	Human protein SEQ
22	150.5	4.9	743	22	AAW79738	Human protein SEQ
23	149.5	4.9	1353	22	AAW79520	Human protein SEQ
24	149.5	4.9	1367	22	AAW78536	Novel human diago
25	149.5	4.9	1356	22	ABG22570	Novel human diago
26	148.5	4.9	1396	22	ABG22574	Novel human diago
27	147	4.8	1408	22	ABG22573	Growth factor rece
28	146.5	4.8	217	18	AAW18063	Modified human/mou
29	146	4.8	1567	23	ABH08236	Modified human/mou
30	146	4.8	1594	23	ABH08237	Human ShcC mutant
31	144.5	4.7	474	18	AAW39089	Human ShcC mutant
32	144.5	4.7	474	18	AAW39087	Human ShcC mutant
33	142.5	4.7	474	18	AAW39084	Human ShcC mutant
34	142.5	4.7	474	18	AAW39092	Human ShcC mutant
35	141.5	4.6	474	18	AAW39085	Human ShcC mutant
36	141.5	4.6	474	18	AAW39088	Human ShcC mutant
37	141.5	4.6	474	18	AAW39090	Human ShcC mutant
38	141.5	4.6	474	18	AAW39091	Human ShcC mutant
39	141.5	4.6	474	18	AAW39095	Human ShcC mutant
40	141.5	4.6	474	18	AAW39096	Human ShcC mutant
41	141.5	4.6	474	18	AAW23246	Human ShcC protein
42	141.5	4.6	474	18	AAW15257	Rat brain-specific
43	141	4.6	728	22	AAW18754	Human protein SEQ
44	140.5	4.6	469	17	AAW97242	SCK phosphotyrosin
45	140.5	4.6	474	18	AAW39094	Human ShcC mutant

ALIGNMENTS

RESULT 1

AA881070 standard; Protein: 435 AA.

AA881070;

25-JUN-2001 (first entry)

Murine mast cell-specific signal transduction protein.

Mast cell; signal transduction; mouse; allergic disease.

Mus musculus.

JP3146204-B1.

12-MAR-2001.

17-SEP-1999; 99JP-0263778.

XX 17-SEP-1999: 99JP-0263778.
 XX (KAGA-) KAGAKU GIUTSU SHINKO JIGYODAN.
 PA WPI: 2001-310022/23.
 DX N-PSDB: AAF86139.
 DR
 XX
 PT Mast cell-specific signal-transduction molecule, useful for screening
 PT therapeutic compounds for treating allergies, is specifically expressed
 PT by mouse mast cell.
 XX
 PS Claim 1: Page 7-8; 12pp: Japanese.
 XX
 CC This sequence represents a murine mast cell-specific signal transduction
 CC protein. The invention includes the cDNA and protein sequences of the
 CC mast cell-specific signal transduction molecule and an expression vector
 CC containing the polynucleotide sequence. The coding sequence of the signal
 CC transduction protein can be used for screening therapeutic compounds
 CC which will be useful for treating allergic diseases.
 XX
 SQ Sequence 435 AA:

Alignment Scores: 1.2e-203 Length: 435
 Pred. No.: 2316.00 Matches: 435
 Score: 2316.00 Conservative: 0
 Percent Similarity: 100.00% Mismatches: 0
 Best Local Similarity: 100.00% Indels: 0
 Query Match: 75.76% Gaps: 0
 Db: 22

US-09-856-061-1 (1-1721) x AAB81070 (1-435)

OY 255 ATGACACGACGAGGCAATATAAGACAAAGAAAGATGCGTGATGATTCGAC 314
 DB 1 MetThrSerGlnGlyAsnLysArgThrThrLysGluGlyPheGlyAspLeuArgPheGln 20
 OY 315 AACGTCCTCTGCTGAAAAATAGTCATGCGCAAGCCTCAGACAGTCCCAAGGGCGGTG 374
 DB 21 AsnValSerLeuLeuLysAsnArgSerTrpProSerLeuSerSerAlaLysGlyArgCys 40
 OY 375 CGAGCGGTCCTGGAACACACTCCGGATCAGACAGAAAGATTGGCGGGGTCCTCCAGTGA 434
 DB 41 ArgAlaValLeuGlnProLeuProAspHisAlaArgAsnLeuAlaGlyAlaProGlyGlu 60
 OY 435 GAAAAATGCACACGATPACACAGCAGTACGAAGATCCTGAGTTCCAGCTGCTGAGGACATG 494
 DB 61 GluLysCysAsnSerAsnAsnAspTyrGluAspProGluPheGlnLeuLeuLysAlaTrp 80
 OY 495 CCATCAATGAANAATTTTACAGCCAGACCTATCCAGGATCGGATPACGACATACAGC 554
 DB 81 ProSerMetLysIleLeuProAlaArgProIleGlnGlnSerGlnTyrAlaAspThrArg 100
 OY 555 TATTTCCAGGATATGATGAGGCTCCCTTCTGTTACCTCCCAAGGCTTGTCTCCACT 614
 DB 101 TyrPheGlnAspMetMetGlnAlaProLeuLeuLeuProProLysAlaSerValSerThr 120
 OY 615 GAGACACAAACGAGGATGTGAGATGACACAGCTGGAAGAAGTGACAGACCTTACTTC 674
 DB 121 GlnAArgGlnThrArgAspValAlaArgMetThrGlnLeuGlnGluValAspLysProThrPhe 140
 OY 675 AAGGATGTCAGAAAGCAAGCGTTTAAAGATTCAATACAAAAATAACAAAGACTCTTC 734
 DB 141 LysAspValArgSerGlnArgPheLysGlyPheLysTyrThrLysIleAsnLysThrPro 160
 OY 735 TTGCCACCTCTCGGCTGCTACTACTCTCCCAAGAGATGACCAACCTTACCCCCACAGA 794
 DB 161 LeuProProProArgProAlaIleThrLeuProLysLysTyrGlnProLeuProProAla 180
 OY 795 CCACACAGAGGAGAGCAGTACTTCTGCTCCAAAGCCACCTTCCAGAAAGTCCAGAG 854
 DB 181 ProProGlnGlnSerSerAlaTyrPheAlaProLysProThrPheProGluValGlnArg 200

OY 855 GGGCCAGGCGAGAGAGTGCAAAAAGACTTCAGTAGGGTCTTGAGACAGAAAGAAATCT 914
 DB 201 GlyProArgGlnArgSerAlaLysAspPheSerArgValLeuGlnGluGlnSer 220
 OY 915 CACCACACAGCAAGACGACAAATCTTTGCCCCATCTCAACCAACCAACACAGAAAGT 974
 DB 221 HisHisGlnThrLysProGlnSerSerCysProSerSerAsnGlnAsnThrGlnLysSer 240
 OY 975 CCACCTGCCATTGCCAGCTCTTCTTCATGCGCAGAAACACAGTATACAGCAGAGAC 1034
 DB 241 ProProAlaIleAlaSerSerSerTyrMetProGlyLysHisSerIleGlnAlaArgAsp 260
 OY 1035 CATACAGGTACATGACAGACTGTCTGCTCAGAGATGCCAAGCTGACAGCCACAGC 1094
 DB 261 HisThrGlySerMetGlnHisCysProAlaGlnArgCysGlnAlaAlaAlaSerHisSer 280
 OY 1095 CCGCAATGCTCGCCCATGAAAAACAACTGGGAAACCTGACCCCAAGCTGTAT 1154
 DB 281 ProArgMetLeuProTyrGlnAsnThrAsnSerGluLysProAspProThrLysProAsp 300
 OY 1155 GAGAAAGATGCTGCGCAGAAATGATGATGATGAGATGAGATGACAGTCGACGAGTGA 1214
 DB 301 GluLysAspValTrpGlnAsnGlnLyrTrpTyrIleGlyGluTyrSerArgGlnAlaValGlu 320
 OY 1215 GATGTCTTAATGAAGAGACAGAGATGCTATTTTGGTCCGAGACTGCTTACAAAA 1274
 DB 321 AspValLeuMetLysGlnLysAsnLysAspGlyThrPheLeuValAlaArgAspCysSerThrLys 340
 OY 1275 TCCAAAGCGCAACCATATGTTGTTGGTGTGTTTATGGAACAGAGTCTACAAATGAAA 1334
 DB 341 SerLysAlaGlnProTyrValLeuValAlaPheTyrGlyAsnLysValTyrAsnAllys 360
 OY 1335 ATCCGTTTCTCTGAGACCAATACACAGTTTGGCCCTGGCACAGACTGCTACAAAG 1394
 DB 361 IleArgPheLeuGlnSerAsnGlnInPheAlaLeuGlyThrGlyLeuArgGlyAsnGlu 380
 OY 1395 ATGTTGATTTCTGTGGAAGACATCATTAACACACTACATATTTCCCATCTGCTATA 1454
 DB 381 MetPheAspSerValGlnAspIleIleGlnHisTyrThrTyrPheProIleLeuLeuIle 400
 OY 1455 GATGGGAAAGACAGAGCTGACGACGAGAAACAGTGTACTACACAGCACTGCTTC 1514
 DB 401 AspGlyLysAspLysAlaAlaAlaArgArgLysGlnCysTyrLeuThrGlnProLeuProLeu 420
 OY 1515 GCCAGGCTCTCTTCATCAGTACAGTCCAGCCAGGACACTTCATGAG 1559
 DB 421 AlaArgLeuLeuLeuThrGlnTyrSerSerGlnAlaLeuHisGlu 435

RESULT 2
 AAE22608
 ID AAE22608 standard; Protein: 443 AA.
 XX
 AC AAE22608;
 XX
 DT 26-JUL-2002 (first entry)
 XX
 DE Human MIST protein #1.
 XX
 KW Human: mast cell immunoreceptor signal transducer; MIST; Immune disorder;
 KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
 KW inflammation; cancer; gene therapy; cyostatic; immunomodulatory.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 84
 FT Modified-site 111
 FT Modified-site 306..311
 FT Domain /label= Tyrosine_phosphorylation_binding_site
 FT Domain /label= SH3_binding_proline-rich_motif
 FT Domain 324..407
 FT /label= SH2_domain

```

XX WO200226986-A2.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US30593.
XX
XX 29-SEP-2000; 2000US-237030P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX
XX WPT. 2002-372126/40.
XX
XX N-PSDB; AAD35800.
XX
XX New isolated mast cell immunoreceptor signal transducer polypeptide,
XX useful for treating immune disorder involving hyperactivity of B- or
XX T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
XX
XX Claim 11; Fig 3; 171pp; English.
XX
XX The present invention relates to novel mast cell immunoreceptor signal
XX transducer (MIST) proteins and polynucleotides encoding such proteins.
XX MIST sequences of the invention are useful for preventing, treating or
XX ameliorating a medical condition in mammalian subject. They are useful
XX for treating an immune disorder involving hyperactivity of B- or T-
XX lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX lymphoma, tumor or thymoma in a mammal. MIST sequences are useful
XX as targets for therapeutic intervention in immune cell disorders and
XX inflammatory indications, for diagnosis and/or screening of disorders
XX or diseases associated with expression of MIST, for screening for
XX antagonists or inhibitors of the interaction of MIST with cellular
XX signaling components. They are used in assays that detect activation
XX or induction of various B and T-cell-related neoplasms or cancers.
XX Sequences of the invention are also used in gene therapy. The present
XX sequence is human MIST protein.
XX
XX Sequence 443 AA:
XX
XX Alignment Scores:
XX Pred. No.: 7,59e-117 Length: 443
XX Score: 1370.50 Matches: 278
XX Percent Similarity: 73.30% Conservative: 46
XX Best Local Similarity: 62.90% Mismatches: 113
XX Query Match: 44.83% Indels: 5
XX DB: 23 Gaps: 4
XX
XX US-09-856-061-1 (1-1721) x AAE22608 (1-443)
XX
XX 213 GCTGAGTTGAAGATCCCTCTTTCGTCAGTCCAGAGCATGACACAGGCGCAT 272
XX |||||||
XX DB 2 AAGTleuylsileprleuthrArglnValProargThrMetAsnArgGlnGlyAsn 21
XX
XX 273 AAAAGCAGCAAGAAAGAGATTGCGTATGATTCAGATTCAGAAAGCTCTCTGTA 332
XX |||||||
XX DB 22 ArglyshrhrlrlysglnGlySerAsnAspLeuylsPheGlnAsnPheserLeuProlys 41
XX
XX 333 AATAGTCATAGCCAGCTCAGCAGTGCAGAAAGCGCGGTGTCAGCGGTTTGGAA 392
XX |||||||
XX DB 42 AsnArgserTrpProArgIleAsnSerAlaThrGlyInlryGlnArgMetAsnLysPro 61
XX
XX 393 CTTCGGATCAGCAAGAACTTGTGCGGTGCTCCAGTCCAGTGGAGAAATGCAAGTA 452
XX |||||||
XX DB 62 LeuLeuAspTrpGlnValArgAsnPhelAlaValAlleuAspGlyAlaLysGlyHisSerAsp 81
XX
XX 453 AACCACTAGCAAGATCCCTGATTCAGCTGCTGAGGAGCCATCATGAAATGTA 512
XX |||||||
XX DB 82 AspAspTrpIleAspAspProGlnLeuArgMetGlnGlnTrpTrpGlnInsrlleuLysleu 101
XX
XX 513 CCAGCCAGACTATTCAGGAATGCGAATACGCAATACAGCTATTTTCAGAGATATG 572
XX |||||||
XX DB 102 ProAlaArgProIleLysGlnSerGlnValArgLysPheHisTyrPheLysValAlaMet 121

```

```

OY 573 GAGGCTCCCTCTGTGTACTCTCCAGGCTTCTCTCTCAGTGAAGACAACAGGAT 632
OY |||||
DB 122 AspIleProleuProleuAspThrArgThrSerIleSerIleGlyGlnProThrTrpAsn 141
OY 633 GTGAGGATGACACAGCTGGAAGAGTGAACAGCTTACCTTCAGAGATGCAAGCA 692
OY |||||
DB 142 ThrGln---ThrArgLeuGlnValArgLysProIleSerLysAspValArgSerGln 160
OY 693 CGCTTAAAGATTTCAATATACCAAAATTAACAGAGCTCTTTCGCCACCTTCGGGCT 752
OY |||||
DB 161 AsnIleLysGlyAspAlaSerValArgLysAsnLysIleProleuProProArgPro 180
OY 753 GCTATCAGTCCCTCCAGAGATGACACAGCTTACCCAGCAGCAGAGAGAGCAGT 812
OY |||||
DB 181 LeuIleThrLeuProLysLysTrpGlnProleuProPro---GlnProGlnSerArg 199
OY 813 GCATACCTGCTCCAAAGCCACCTTTCAGAGATGTCAGAGGGGCCAGCAGAGAGT 872
OY |||||
DB 200 ProProleuSerGlnArgIleThrPheProGlnValGlnArgMetProSerGlnIleSer 219
OY 873 GCAAAAGACTTACGTAGGCTCCTTGAGCAGAAAGAAATCTCACACACAGCAAGCA 932
OY |||||
DB 220 LeuArgAspLeuSerGlnValleuGlnAlaGlnLysValProHisAsnGlnArgLysPro 239
OY 933 GAATCTTCTGCCCATCATCAAAACCAACACAGCAAGAGTCCACCTCCATTGCGAGC 992
OY |||||
DB 240 GluSerThrHisLeuLeuGlnAsnGlnAsnThrGlnGlnIleProleuAlaIleSer 259
OY 993 TCTTCTATACGCGCAGAAAGACAGTATACAGCAGAGACCAATACAGTATGATG 1052
OY |||||
DB 260 SerSerPheThrThrSerAsnHisSerValGlnAsnArgAspHisArgLysGlyMetGln 279
OY 1053 CACTGCTCTGCTCAGAGATGCCAGCTGCAGCCAGCAGCAAGCCCT-----CGAATGCTG 1106
OY |||||
DB 280 ProCysSerProGlnArgGlySerGlnProAlaSerCysSerProHisGlnAsnIleLeu 299
OY 1107 CCTATGAAACACAACTCGAGAAACCTGAGAAACCTGAGCAAGCTGATGAAAGATGTC 1166
OY |||||
DB 300 ProTyrLysIleThrSerTrpArgProProPheProLysArgSerAspArgVal 319
OY 1167 TGGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1226
OY |||||
DB 320 GlnHisAsnGlnIleProLysIleGlyLysTrpArgGlnAlaValGlnAlaPheMet 339
OY 1227 AAAGAGAACAGAGATGCTATTTTGGTCCGAGATGCTCTCAAAATCCAGAGCA 1286
OY |||||
DB 340 LysGlnAsnLysAspGlySerPheLeuValArgAspCysSerThrLysSerLysGlnGly 359
OY 1287 CCATATGTTTGGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1346
OY |||||
DB 360 ProTyrValleuAlaValPheTyrGlnAsnLysValIleValLysIleArgPheLeu 379
OY 1347 GAGAGCAATCAACAGTGGCTGCGGACAGCAAGCTACAGAGAAATGATGTTGAT 1406
OY |||||
DB 380 GlnArgAsnGlnGlnPheAlaLeuGlyThrGlyLeuArgGlyLysProLysPheAspSer 399
OY 1407 GTGGAACATCAATGTAACATCAACATCAATGTTCCATGTTGTAATGATGGAAG 1466
OY |||||
DB 400 ValGlnAspIleIleGlnIleTyrLysAsnPheProIleLeuIleAspGlyLysAsp 419
OY 1467 AAG---GCTCAGCAGCAAGAAAGTGTCTACCTCAGCAGCAGCTGCTGCGAGGCTC 1523
OY |||||
DB 420 LysThrGlyValHisArgLysGlnCysHisLeuThrGlnProleuProleuThrArgHis 439
OY 1524 CTCTCTC 1529
OY |||||
DB 440 LeuLeu 441
OY |||||

```

RESULT 3
 AAE22609
 ID AAE22609 standard; Protein: 428 AA.
 XX


```

DB      419  LeuProLeuThrArgHisLeuLeu 426
|||||
RESULT 4
AAB81071
ID      AAB81071 standard; Protein: 376 AA.
XX
XX      AAB81071;
XX
XX      25-JUN-2001 (first entry)
XX
XX      Human mast cell-specific immunoreceptor signal transducer.
XX
XX      Mast cell; signal transduction; human; allergic disease; MIST;
XX      mast cell-specific immunoreceptor signal transducer.
XX
XX      Homo sapiens.
XX
XX      JP3146204-B1.
XX
XX      12-MAR-2001.
XX
XX      17-SEP-1999; 99JP-0263778.
XX
XX      17-SEP-1999; 99JP-0263778.
XX
XX      (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
XX      WPI: 2001-310022/33.
XX      N-PSDB; AAF6140.
XX
XX      Mast cell-specific signal-transduction molecule, useful for screening
XX      therapeutic compounds for treating allergies, is specifically expressed
XX      by mouse mast cell.
XX
XX      Example 1; Page 10-11; 12pp; Japanese.
XX
XX      This sequence represents human MIST (mast cell-specific immunoreceptor
XX      signal transducer). The invention relates to cDNA encoding a murine mast
XX      cell-specific signal transduction protein. Included in the invention are
XX      cDNA and protein sequences of the mast cell-specific signal transduction
XX      molecule and an expression vector containing the polynucleotide sequence.
XX      The coding sequence of the signal transduction protein can be used for
XX      screening therapeutic compounds which will be useful for treating
XX      allergic diseases.
XX
XX      Sequence 376 AA;
XX
XX      Alignment Scores:
XX      Pred. No.: 5.53e-95 Length: 376
XX      Score: 1132.00 Matches: 229
XX      Percent Similarity: 72.07% Conservative: 42
XX      Best Local Similarity: 60.90% Mismatches: 101
XX      Query Match: 37.03% Indels: 4
XX      DB: 22 Gaps: 3
XX
US-09-856-061-1 (1-1721) x AAB81071 (1-376)
OY      309 TTCCAGAACGTCCTCTCTCTGAAATAATGTCATGCGCCACAGCTCAGAGTCCCAAGG 368
|||
DB      1 PheGlnAsnPheSerLeuProLysAsnArgSerTrpProAlaGlnAsnSerAlaThrIly 20
|||
OY      369 CGGTGCGAGCGGTCTCGAAGCACTTCGATCAGACAGAACTTGCTGGTGGGTCCTCCA 428
|||
DB      21 GlnTyrGlnArgMetAsnLysProLeuLeuAspTrpGlnArgAsnPheAlaIleValLeu 40
|||
OY      429 GGTGAGAAATGCAACAGTACAGACAGTACGAAATCTGAGTTCACGCTGCTGAAG 488
|||
DB      41 AspGlyAlaLysGlyHisSerAspAspIlyTrpAspAspProGlnLeuArgMetLysGln 60
|||
OY      489 GCATGGCCATCAATGAAATTTTACAGCCAGACTATCCAGAGTTCGGAATACGACAGT 548
|||
DB      61 ThrTrpGlnSerIleLysIleuProAlaArgProIleLysGlnSerGlnTyrAlaAsp 80

```

```

OY      549 ACACGCTATTTCAGAGATATGATGAGGCTCCCTTCGTGTACTCCACAGGCTTCTGTC 608
|||
DB      81 ThrHisTyrPheLysValAlaIleLasPheProLeuPheLeuAspThrArgHisSerIle 100
|||
OY      609 TCCACTGAGAGACAAACAGGAGTGTGAGATGACACAGCTTGGAAGATGGACAAGCT 668
|||
DB      101 SerIleGlyGlnProThrTrpAsnThrGln---ThrArgLeuGlnArgValAspLysPro 119
|||
OY      669 ACCCTTCAGAGATGTCAGAAAGCCACGCTTTAAAGATTCMAATACCAAAATTAACAAG 728
|||
DB      120 IleSerArgAspValArgSerGlnAsnIleLysGlyAspAlaSerValArgLysAsnLys 139
|||
OY      729 ACTGCTTGCCACCTCCCTGGGCTGCTATCACCTCCGCCAAGATACCAACCTTACCC 788
|||
DB      140 IleProLeuProProArgProLeuIleThrLeuProLysLysTyrGlnProLeuPro 159
|||
OY      789 CCAGCACACACAGAGAGAGAGAGTGCATCTTCGCTCCAAAGCCACCTTCCAGAGTC 848
|||
DB      160 Pro---GlnProGlnSerSerArgProProLeuSerGlnArgHisThrPheProGlnVal 178
|||
OY      849 CAGAGGGGGCCAGGCGACAGAGAGTGCAAAAGCTTACAGTAGGCTCCTTGACAGAGAA 908
|||
DB      179 GlnGlyMetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnArgLys 198
|||
OY      909 GAATCTCACCCACAGACAAAGCCAGAACTCTTCGCCATCAACCAACCAACACACAG 968
|||
DB      199 ValProHisAsnGlnArgLysProGlnSerThrHisLeuLeuGlnAsnGlnAsnThrGln 218
|||
OY      969 AAGAGTCCACCTCCATTCGACCTCTTCCTACATCCAGAAAGACAGATATACAGCC 1028
|||
DB      219 GlnIleProLeuAlaIleSerSerSerPheThrThrSerAsnHisSerValGlnAsn 238
|||
OY      1029 AGAGACCATACAGTACATGACGACACTGCCCTGCCATCAGAGATGCCAGCTGACGACG 1088
|||
DB      239 ArgAspHisArgGlyGlyMetGlnProCysSerProGlnArgGlyGlnProProAlaSer 258
|||
OY      1089 CACAGCCCT-----CGAATGCTGCCCTATGAAACAAACAACTCGAGAAACCTGACCC 1142
|||
DB      259 CysSerProHisGlnAsnIleLeuProTyrLysTyrThrSerTrpArgProProPhePro 278
|||
OY      1143 ACAAGCCCTATGAGAGATGTCGTCGACAAATGATGTACATTTGGAGAAATACACTCCG 1202
|||
DB      279 LysArgSerAspArgLysAspValGlnHisAsnGlnIlyTrpIleGlyGlnTyrSerArg 298
|||
OY      1203 CAGGCACTGGAAATGCTTAAATGAAGAGAAACAGAGTACTTTTGTGCTCCGAGAC 1262
|||
DB      299 GlnAlaValGlnGlnAlaPheMetLysGlnAsnLysAspLysSerPheLeuValArgAsp 318
|||
OY      1263 TGCCTACAAATTCAGAGCAGAAACCATATGTTTGGTGTGTTATGGCAACAGGTC 1322
|||
DB      319 CysSerThrLysSerLysGlnGlnProTyrValLeuAlaValPheTyrGlnAsnLysVal 338
|||
OY      1323 TACAATGTGAAATCCCTTCCTCCAGAGCAATCAACAGTTCGCTCCGACAGAGCTA 1382
|||
DB      339 TyrAsnValLysIleArgPheLeuGlnArgAsnGlnIlePheAlaLeuGlyThrGlyLeu 358
|||
OY      1383 CGAGGAATGAGATGTTGATTTCTGTGGAAGACATCATTTGAACATAC 1430
|||
DB      359 ArgGlyAspGlnLysPheAspSerValGlnLysPheIleIleGlnHisTyr 374
|||
RESULT 5
AAE22610
ID      AAE22610 standard; Protein: 353 AA.
XX
XX      AAE22610;
XX
XX      26-JUL-2002 (first entry)
XX
XX      Human MIST splice variant protein from clone #12.
XX
XX      Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
XX      lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;

```

KM Inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
 OS Homo sapiens.
 PN WO200226986-A2.
 XX
 XX 04-APR-2002.
 PD
 PF 28-SEP-2001; 2001WO-US30593.
 XX
 PR 29-SEP-2000; 2000US-237030P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
 PI MPI: 2002-372126/40.
 DR N-PSDB: AAD35802.
 XX
 XX
 PT New isolated mast cell immunoreceptor signal transducer polypeptide,
 PT useful for treating immune disorder involving hyperactivity of B- or
 XX T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
 XX
 PS Claim 11, Fig 8; 171pp; English.
 XX

CC The present invention relates to novel mast cell immunoreceptor signal
 CC transducer (MIST) proteins and polynucleotides encoding such proteins.
 CC MIST sequences of the invention are useful for preventing, treating or
 CC ameliorating a medical condition in mammalian subject. They are useful
 CC for treating an immune disorder involving hyperactivity of B- or T-
 CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
 CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
 CC as targets for therapeutic intervention in immune cell disorders and
 CC inflammatory indications, for diagnosis and/or screening of disorders
 CC or diseases associated with expression of MIST, for screening for
 CC antagonists or inhibitors of the interaction of MIST with cellular
 CC signalling components. They are used in assays that detect activation
 CC or induction of various B and T-cell-related neoplasms or cancers.
 CC Sequences of the invention are also used in gene therapy. The present
 CC sequence is human MIST splice variant protein from clone #12.
 CC

50 Sequence 353 AA;

Alignment Scores:

Pred. No.:	6,25e-03	Length:	353
Score:	1109.50	Matches:	226
Percent Similarity:	73.22%	Conservative:	31
Best Local Similarity:	64.39%	Mismatches:	89
Query Match:	36.29%	Indels:	5
DB:	23	Gaps:	4

US-09-856-061-1 (1-1721) x AAE22610 (1-353)

[illegible]

QY	786	CCCCAGACACACACAGAGAGAGAGATGATCTTCTGGCTCCAAAGCCACCTTCCAGAA	843
Db	102	ProPro---GluProGluSerSerArgProProLeuSerGlnAlaSerThrPheProGlu	120
QY	846	GTCCAGAGGGGGCCAGGACAGAGAGTGCMAAAGACTTCAGTAGAGCTCTGGAGCAAA	905
Db	121	ValGlnAlaGlyeProSerGlnIleLeuSerLeuArgGAspLeuSerGluValIleGlnAlaGlu	140
QY	906	GAAGAAATTCACACACACAGCAAAAGCCAGAAATCTTCTCCATATCAAAACCAACACA	965
Db	141	LysValAlaProHisAsnIleArgGlyAspGluProGluSerThrIleHisLeuGluAsnGlnAsnThr	160
QY	966	CAGAAAGATCCACACTCCATTCGACAGCTCTTCATCATATGCCAGCAAAAGCAGATATCAAA	102
Db	161	GlnGlnIleProLeuAlaIleLeuSerSerSerPheThrIleHisAsnIleSerValGln	180
QY	1026	GCCAGAGACCATATACAGTATAGCATGCAGACATGCTCCCTGCTCAGAGATGCCAAGCTGCAGCC	108
Db	181	AsnAlaArgAspHisArgGlyGlyMetGlnProCysSerProGlnArgCysGlnProProAla	200
QY	1086	AGCCACAGCCCT-----CGAATGGTGGCCATATGAAACCAAACTGGGAAACCTGCAC	113
Db	201	SerCysSerProHisGlnAlaAsnIleLeuProTyrIleSerThrSerTrpArgProProPhe	220
QY	1140	CCGCAAAAGCCTGATGAGAGAGATCTCTGCAGAGATGAATGATATTCAGTGGAGAAATACAGT	119
Db	221	ProLysArgSerAspArgGlyAspArgAlaGlnHisAsnGlnIleProTyrIleGlyGlyIleTyrSer	240
QY	1200	CGCGAGGCGAGTGGAGAGATGTGTATATGAAGAACAACAAGATGGTACTTTTGGTCCGA	125
Db	241	ArgGlnAlaValAlaGlnAlaIlePheMetCysGlnAlaAspIleLysAspIleSerPheLeuValArg	260
QY	1260	GACTGCTCTACAAATCCCAAGGCAACCAATATGTGTTGGTGGTATTATGGAGCAAG	131
Db	261	AspCysSerThrLysSerLysGlnIleProTyrAlaIleuAlaValAlaPheTyrGlnAlaSer	280
QY	1320	GTCTACATGTGGAAATCCGTTTCTCTGCAGAGCATACACAGTTTCCCTGGGACACAGA	137
Db	281	ValTyrAsnValLysIleArgPheLeuGlnArgAsnGlnIlePheAlaLeuGlyThrGly	300
QY	1380	CTACAGCAAAATAGATGTGATTCGTGTGAAGACATCATATGAACATACACATATTTT	143
Db	301	LeuAlaGlyAspArgGlnLysPheAspSerValGlnAspIleIleGlnIleTyrLysAsnIle	320
QY	1440	CCCATCTGTGTAATAGATGGGAAAGACAAAG--GCTGCAGCGAGCAAGAAACAGTCTACCTC	1498
Db	321	ProIleIleLeuIleAspArgLysAspLysTyrGlyValHisArgGlySerGlnGlnHisLeu	340
QY	1497	ACCAGGCCACGCTCTCGCAGGCTCCTTCTC	1529
Db	341	ThrGlnProLeuProLeuThrArgHisLeuLeu	351
RESULT 6			
ID	AAB93406	standard; Protein; 456 AA.	
XX	AAB93406;		
AC	26-JUN-2001	(first entry)	
DT	26-JUN-2001	(first entry)	
XX	Human protein sequence seq	ID NO:12602.	
DE	Human protein sequence seq	ID NO:12602.	
XX	Human; primer; detection; diagnosis; antisense therapy; gene therapy.		
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy.		
XX	Homo sapiens.		
OS	Homo sapiens.		
XX	EpI074617-A2.		
PN	07-FEB-2001.		
XX	07-FEB-2001.		
PD	28-JUL-2000.	2000EP-0116126.	
PF	28-JUL-1999.	99JP-0248036.	
XX	28-JUL-1999.	99JP-0248036.	
RR	28-JUL-1999.	99JP-0248036.	

RESULT 6	
AAB93406	
ID	AAB93406 standard; Protein; 456 AA.
XX	
XX	
AC	AAB93406;
XX	
DT	26-JUN-2001 (first entry)
XX	
DE	Human protein sequence seq ID NO:12602.
XX	
KW	Human; primer; detection; diagnosis; antisense therapy; gene therapy.
XX	
XX	
OS	Homo sapiens.
XX	
PN	EP1074617-A2.
XX	
PD	07-FEB-2001.
XX	
PE	28-JUL-2000; 2000EP-0116126.
XX	
DR	29-JUL-1999; 99JP-0248036.

PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 PA (HELI-) HELIX RES INST.
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 DR WPI: 2001-318749/34.
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS
 PS Claim 8: SEQ ID 12602; 2537bp + CD ROW; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to
 CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 XX Sequence 456 AA:
 SQ
 Alignment Scores:
 Pred. No.: 5 91e-21 Length: 456
 Score: 326.00 Matches: 99
 Percent Similarity: 37.06% Conservative: 50
 Best Local Similarity: 24.63% Mismatches: 141
 Query Match: 10.66% Indels: 112
 DB: 22 Gaps: 9
 US-09-856-061-1 (1-1721) x AAB93406 (1-456)
 QY 444 AACAGTACACGACTAGAGAGCTGAGTTCAGCTGCGAAGCGATGCCATCATG 503
 DB 91 AsnAlaAspSerPheGluProProProValGlu-----GlnGluThrArg 106
 QY 504 AAAATTTTACGACGACATCATCAGATCGAATACGATACAGCTATTTCAG 563
 DB 107 ProValHisProAlaLeuProPheAlaArgGluGluTyrIleAspAsnArgSerSerGln 126
 QY 564 GATATGATGAGGCTCC-----CTTCTGTTACTCCCAAGGCTTGTCTCCACT 614
 DB 127 -----ArgHisSerProProPheSerIleThrLeuProSerIleProSer 144
 QY 615 GAGAGACAAACAGGATGTGAGATGACACAGCTGGAAGAAGTGACACGCTACCTTC 674
 DB 145 GluLysAlaArgLeuThrSerThrLeuProAlaLeuThrAlaLeuGlnLysProGlnVal 164
 QY 675 AAGGATGTACAGACCAACGCTTTAAAGATTCAAATTCACAAAATAACAGACTGCT 734
 DB 164 ----- 164

QY 735 TTGCACCTCTCGGCGCTGATATCATCTCCCAAGAGTACCAACCTTACCCACAGA 794
 DB 165 -----ProProLysProLysGlyLeuLeuGlnLysGlnAlaAspTyrValValProVal 182
 QY 795 CCACACGAGAGAGAGAGTGCATCTTCCCTCCCAAGCCACCTTTCAGAAAGTCCAGAG 854
 DB 183 GluAspAsnAspGluAsnTyrIleHisProThrGlnSerSerProProGluLys 202
 QY 855 GGGCCCGAGGACAGAGATGTCAAAGACTTCAGTAGGGTCTTTGGACGAAAGAGATCT 914
 DB 203 AlaProMet-----Val 206
 QY 915 CACCAACAGACAAAGCCAGAAATCTTGTCCCATCATCAAAACCAACACA----- 965
 DB 207 AsnArgSerThrLysProAsnSerSerThrProAlaSerProProGluThrAlaSerGly 226
 QY 966 -----CAGAAGAGTCCAGCTGCCATTGCCAGCTCT----- 995
 DB 227 ArgAsnSerGlyAlaThrPgluThrLysSerProProAlaAlaProSerProLeuPro 246
 QY 995 ----- 995
 DB 247 ArgAlaGlyLysLysProThrThrProLeuLysThrThrProValAlaSerGlnLysn 266
 QY 996 -----TCTACATGCGAAGAAAGACAGTATACAAAGCAGACCATACAGGTAGCATGCG 1052
 DB 267 AlaSerSerValCysGlyLysLysProLysProLysProLysGlnAlaGlnArgHisArgLysSerHis 286
 QY 1053 CACTGTCTGCTCAGAGATCCCAAGCTGCAGCCAGCCAGCCCTGGAATGCTGCCCTAT 1112
 DB 287 -----ArgGlnGlnAlaValGlnSerProValPheProProAla 299
 QY 1113 GAAACACAAACCTCGGAGAACTGACCC----- 1142
 DB 300 GlnLysGlnIleHisGlnLysProLysProLysProLysProLysPheThrGlnLysnPro 319
 QY 1143 -----ACAAAGCCTGATGAGAG 1160
 DB 320 ThrValAspLysProLeuProSerPheSerSerAsnSerThrIleSerGlnGlnAla 339
 QY 1161 GATGCTGCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1220
 DB 340 GlyValLeuLysLysProPyrAlaGlyAlaCysAspArgLysSerAlaGluGluAla 359
 QY 1221 TTAATGAAGAGACAGAGATGCTACTTTTGTGCTCCAGACTGCTCAAAATCCAG 1280
 DB 360 LeuHisArgSerAsnLysAspLysSerPheLeuIleArgLysSerSerGlnHisAspSer 379
 QY 1281 GCAGAACCATATGTTGT 1340
 DB 380 LysGlnProTyrThrLeuValAlaPhePheAsnLysArgValLysIleProValArg 399
 QY 1341 TTCCTGAGAGCATCAACAGTTTGCCTGCGACAGACTAGACGAATGATGATG 1400
 DB 400 PheIleGluAlaThrLysGlnTyrAlaLeuGlnArgLysLysAsnGlyGluGluTyrPhe 419
 QY 1401 GATCTGTGGAAGACATCATGAACTACATCACTATATTTCCATTCGTATAGATGG 1460
 DB 420 GlySerValAlaGluIleIleArgAsnHisGlnHisSerProLeuValLeuIleAspSer 439
 QY 1461 AAAGAC 1466
 DB 440 GlnAsn 441
 RESULT 7
 AAM59866 standard; Protein: 449 AA.
 XX
 AC AAM59866;
 XX
 DT 20-NOV-1998 (first entry)

KW ras pathway; GDP; GTP; calcium pathway; antigen.
 OS Homo sapiens.
 XX Key Location/Qualifiers
 FH MISC-difference 327
 FT /note= "encoded by AGC"
 XX
 XX WO9832852-A1.
 XX
 XX 30-JUL-1998.
 XX
 XX 23-JAN-1998; 98MO-US01394.
 XX
 XX 17-MAR-1997; 97US-0819013.
 XX
 XX 24-JAN-1997; 97US-0788322.
 XX
 XX (UNIV) UNIV WASHINGTON.
 XX
 XX Chan AC, Fu C;
 XX
 XX WPI; 1998-427948/36.
 DR N-PSDB; AAV41901.
 XX
 XX Human B cell linker proteins - useful in the treatment of diseases
 PT Involving increased or decreased apoptosis
 XX
 XX Claim 4; Fig 1; 56pp; English.
 XX
 XX This is the amino acid sequence of the human BLNK-1 (B cells linker
 CC protein-1) protein, used the treatment of diseases involving the
 CC increase and decrease of apoptosis, in the method of the invention.
 CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
 CC intracellular calcium levels and Grb2 is an adapter molecule
 CC containing two SH3 domains that mediate its interaction with the
 CC guanine nucleotide exchange factor, Son of Sevenless (SOS) which in
 CC turn activates the ras pathway by facilitating the exchange of GDP for
 CC GTP on the ras molecule. Activation of both ras and calcium pathways
 CC are required for efficient B cell antigen receptor function. BLNK
 CC binding proteins can be used to identify BLNK proteins in a target
 CC sample.
 CC
 CC Sequence 456 AA;
 SQ
 Alignment Scores:
 Pred. No.: 2.6e-20 Length: 456
 Score: 319.00 Matches: 99
 Percent Similarity: 39.65% Conservative: 58
 Best Local Similarity: 25.00% Mismatches: 139
 Query Match: 10.44% Indels: 100
 DB: 19 Gaps: 12
 US-09-856-061-1 (1-1721) x AAM59865 (1-456)
 QY 444 AACAGTAACAGAGTACAGAGTCCAGTTCAGAGTGTGAGGATGCCATCAATG 503
 DB 91 AsnAlaAspAspSerTyrGluProProValGlu-----GlnGluThrArg 106
 QY 504 AAAATTTTACAGCAGACCTATCCAGAAATGGAAATAGCAGATACAGCTATTTCAG 563
 DB 107 ProValHisProAlaLeuProHeaAlaArgGlyGluTyrIleAspAsnArgSerSerGln 126
 QY 564 GATATGATGAGGCTCC-----CTTCTGTACCTCCCAAGGCTTGTCTTCAGT 614
 DB 127 -----ArgHisSerProProPheSerLysThrLeuProSerLysProSerTrpProSer 144
 QY 615 GAGAGCAAAACAGGAGTGTGAGATGACACAGCTGAGAAAGTGGACAGCTTACCTTC 674
 DB 145 GluLysAlaIArgLeuThrSerThrLeuProAlaLeuThrAlaLeuGlnLysProGlnVal 164
 QY 675 AAGGATGTGAGAGCCAGCGCTTAAAGATTCAAAATACACAAAATAAACAAGACTCCT 734
 DB 164 ----- 164

QY 735 TTGGACACCTCCGGGCGCTGCTATCAGCTCCGCCAGAGATCCACAGCCTTACCCAGCA 794
 DB 165 -----ProProLysProLysGlyLeuLeuGluAspGlnLysPyrValValProVal 182
 QY 795 CCACAGAGGAGAGACAGTGTGATCTTCGTCACAAAGCCACCTTCCAGAAATCCAGAGG 854
 DB 183 GluAspAsnAspGlnAsnTyrIleHisProThrGluSerSerSerProProGlnLys 202
 QY 855 GGGCCCAAGCAGAGAGAGCTGCAAAAAGACTTCAATGAGGCTCTGGAGCAGAGAAGATCT 914
 DB 203 AlaProMet-----Val 206
 QY 915 CACCCACAGCAAAAGCCAGATCTTTCGCCATCATCAAAACCAAAACACA----- 965
 DB 207 AsnArgSerThrLysProAsnSerSerThrProAlaSerProProGlyThrAlaSerGly 226
 QY 966 -----CAGAGAGTCCACCTCCATTCGACAGCTTCCATCATG 1004
 DB 227 ArgAsnSerGlyAlaTrpGluThrLysSerProProPro---AlaAlaProSerProLeu 245
 QY 1005 CCA-----GGAAGCAC-----AGTATACAAAGCCAGAGAC 1034
 DB 246 ProArgAlaGlyLysLysProThrThrProLeuLysThrThrProValAlaSerGlnGln 265
 QY 1035 CATACAGGTAGCATG-----CAGCACTGTCTGCTGACAGATGCCAAGTGC--- 1082
 DB 266 AsnAlaSerSerValCysGluGluLysProLeuProAlaGluAlaGlnHisArgLysSer 285
 QY 1083 -----GCCAGCAGAGCCCTCGAATGCTGCCCTATGAAACACAACTCGAG 1130
 DB 286 HisArgGlnGluAlaValGlnSerProValPheProProAlaGlnLysGlnIleHisGln 305
 QY 1131 AACCTGACCCC----- 1142
 DB 306 LysProLysProLeuProAlaArgPheThrGluGlyLysAsnProThrValAspGlyProLeu 325
 QY 1143 -----ACAAAGCCTGATGAGAGATGCTGTGCGACAGATGAA 1178
 DB 326 ProIlePheSerSerAsnSerThrIleSerGlnGlnGlnAlaGlyValLeuCysLysPro 345
 QY 1179 TGTATCATTTGAGAAATACAGTCCGCCAGGCTGAGAGATGTTTAAATGAAGACAGACAG 1238
 DB 346 TrpTyrAlaGlyAlaCysAspArgLysSerSerAlaGlnGlnAlaLeuHisArgSerAsnLys 365
 QY 1239 GATGTACTTTTGTGTCGCGAGACGCTCTCAAAATCCAGAGCAGAACCATATGTTT 1298
 DB 366 AspGlySerPheLeuIleArgLysSerSerGlyHisAspSerLysGlnProTyrThrLeu 385
 QY 1299 GTGTGTTTATGGCAACAGGCTACAAATGGAATCCGTTCCCTCGAGACAAATCAA 1358
 DB 386 ValValPhePheAsnLysArgValTyrAsnIleProValArgPheIleGlnAlaThrLys 405
 QY 1359 CAGTTTGCCCTGGCGACAGACTACAGAGAAATGAGATGTTGATCTGTGGAACATC 1418
 DB 406 GlnTyrAlaLeuGlyArgLysLysAsnGlyGlnGlyTyrPheGlySerValAlaGluIle 425
 QY 1419 ATTGAACATCACATATTTTCCATTCCTGCTAAATAGATGGGAAGAC 1466
 DB 426 IleArgAsnHisGlnHisSerProLeuValLeuIleAspSerGlnAsn 441
 RESULT 9
 AAM59867
 ID AAM59867 standard; Protein; 457 AA.
 AC AAM59867;
 XX
 XX 20-NOV-1998 (first entry)
 DE
 XX Amino acid sequence of the mouse BLNK protein.
 KW Mouse; BLNK; B cells linker protein; apoptosis; Grb2; PLC-gamma;
 SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; SOS;

Db 21 AsnPhSerLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGlyInTrp 40
 QY 375 CGAGGGGTCTGGAACCACTTCGGATCAGAAAGAACTTGCGGGGTCCAGGTGA 434
 Db 41 GlnArgMetAsnLysProLeuLysPTrpGluArgAsnPheAlaValLeuAspGly 60
 QY 435 GAAATAATGCACAGTACACAGCACTACGAAATCCTGAGTTCACAGCTGCGAAGCATGG 494
 Db 61 AlaLysGlyHIsSerAsp****AspAspProGluLeuArgMetGluGluThrTrp 80
 QY 495 CCATCAATGAAA 506
 Db 81 GlnSerIleLys 84
 RESULT 12
 ABB38147
 ID ABB38147 standard; Peptide: 46 AA.
 AC ABB38147;
 XX
 DT 04-FEB-2002 (first entry)
 DE Peptide #5653 encoded by human foetal liver single exon probe.
 XX
 KW Human; foetal liver; gene expression; single exon nucleic acid probe.
 XX
 OS Homo sapiens.
 XX
 PN WO200157277-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00669.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483447/52.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for
 PT analyzing gene expression in human fetal liver -
 XX
 PS Claim 27; SEQ ID NO 30782; 639pp + sequence listing; English.
 XX
 CC The invention relates to a single exon nucleic acid probe for
 CC measuring human gene expression in a sample derived from human foetal
 CC liver. The single exon nucleic acid probes may be used for predicting,
 CC measuring and displaying gene expression in samples derived from human
 CC foetal liver. The present sequence is a peptide encoded by a single exon
 CC nucleic acid probe of the invention.
 CC Note: The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 46 AA;
 Alignment Scores:
 Pred. No.: 6.14e-11 Length: 46
 Score: 212.00 Matches: 41
 Percent Similarity: 91.30% Conservative: 1
 Best Local Similarity: 89.13% Mismatches: 4
 Query Match: 6.93% Indels: 0
 DB: 22 Gaps: 0

US-09-856-061-1 (1-1721) x ABB38147 (1-46)
 QY 1239 GATGGTACTTTTGGTCGAGACTGCTACAAATCCAGCAACCATATGTTTG 1298
 Db 1 AspGlySerPheLeuValArgAspCysSerThrLysSerLysGluGluProTyrValLeu 20
 QY 1299 GTGGGTTTATGGGACACAGCTCTACAAATGTAATCCGTTTCTCGAGAGCAATCAA 1358
 Db 21 AlaValPheTyrGluAsnLysValTyrAsnValLysIleArgPheLeuGluArgAsnGln 40
 QY 1359 CAGTTTGCCCTGGGCACA 1376
 Db 41 GlnPheAlaLeuGlyThr 46
 RESULT 13
 ABB23354
 ID ABB23354 standard; Protein: 46 AA.
 AC ABB23354;
 XX
 DT 23-JAN-2002 (first entry)
 DE Protein #5353 encoded by probe for measuring heart cell gene expression.
 XX
 KW Human; gene expression; heart; microarray; vascular system;
 KW cardiovascular disease; hypertension; cardiac arrhythmia;
 KW congenital heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157274-A2.
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US00666.
 XX
 PR 04-FEB-2000; 2000US-0180312.
 PR 26-MAY-2000; 2000US-0207456.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-0234687.
 PR 27-SEP-2000; 2000US-0236359.
 PR 04-OCT-2000; 2000GB-0024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488899/53.
 XX
 PT Single exon nucleic acid probes for analyzing gene expression in human
 PT hearts -
 XX
 PS Claim 15; SEQ ID NO 25124; 530pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes for
 CC measuring human gene expression in a sample derived from human heart (see
 CC ABA21533-ABA41305). The present sequence is a protein encoded by one such
 CC probe. The probes may be used for predicting, measuring and displaying
 CC gene expression in samples derived from the human heart via microarray.
 CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 46 AA;
 Alignment Scores:
 Pred. No.: 6.14e-11 Length: 46
 Score: 212.00 Matches: 41

Percent Similarity: 91.30% Conservative: 1
 Best Local Similarity: 89.13% Mismatches: 4
 Query Match: 6.93% Indels: 0
 DB: 22 Gaps: 0
 US-09-856-061-1 (1-1721) x ABB23354 (1-46)

QY 1239 GATGCTACTTTTGGTCCGAGACTGCTTACAAAATCCAGGACAGAACCATATGTTTG 1298
 |||||:|||||
 DB 1 ASpglySerPheLeuValArgAspCysSerThrLysSerLysGluGluProTyrValLeu 20
 QY 1299 GTGGTGTATGGAACAAGCTCTACATGTGAATCCGTTCTCGAGACATCAA 1358
 |||||:|||||
 DB 21 AlavalPheTyrGluAsnLysValTyrAsnValLysIleArgPheLeuGluArgAsnGln 40

QY 1359 CAGTTGCCCTGGGCACA 1376
 |||||:|||||
 DB 41 GlnPheAlaLeuGlyThr 46

RESULT 14

AAM58783
 ID AAM58783 standard; Protein; 46 AA.

XX AC AAM58783;

XX DT 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30888.

XX Human; brain expressed exon; gene expression analysis; probe;

KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX KW epilepsy; cancer.

OS Homo sapiens.

PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JUN-2001; 2001WO-US00667.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483446/52.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PT brains -

XX PS Example 4; SEQ ID NO: 30888; 650bp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC brain. They can be used to measure gene expression in brain cell samples,

XX CC which may enable the diagnosis and improved treatment of nervous system

XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX CC epilepsy and cancers. The present sequence is a protein encoded by one of

XX CC the probes of the invention.

XX SQ Sequence 46 AA;

Alignment Scores: 6.14e-11 Length: 46

Pred. No.: 212.00 Matches: 41

Score: 212.00

Percent Similarity: 91.30% Conservative: 1
 Best Local Similarity: 89.13% Mismatches: 4
 Query Match: 6.93% Indels: 0
 DB: 22 Gaps: 0
 US-09-856-061-1 (1-1721) x AAM58783 (1-46)

QY 1239 GATGCTACTTTTGGTCCGAGACTGCTTACAAAATCCAGGACAGAACCATATGTTTG 1298
 |||||:|||||
 DB 1 ASpglySerPheLeuValArgAspCysSerThrLysSerLysGluGluProTyrValLeu 20
 QY 1299 GTGGTGTATGGAACAAGCTCTACATGTGAATCCGTTCTCGAGACATCAA 1358
 |||||:|||||
 DB 21 AlavalPheTyrGluAsnLysValTyrAsnValLysIleArgPheLeuGluArgAsnGln 40

QY 1359 CAGTTGCCCTGGGCACA 1376
 |||||:|||||
 DB 41 GlnPheAlaLeuGlyThr 46

RESULT 15

AAM71296
 ID AAM71296 standard; Protein; 46 AA.

XX AC AAM71296;

XX DT 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31602.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JUN-2001; 2001WO-US00668.

XX PR 04-FEB-2000; 2000US-0180312.

XX PR 26-MAY-2000; 2000US-0207456.

XX PR 30-JUN-2000; 2000US-0608408.

XX PR 03-AUG-2000; 2000US-0632366.

XX PR 21-SEP-2000; 2000US-0234687.

XX PR 27-SEP-2000; 2000US-0236359.

XX PR 04-OCT-2000; 2000GB-0024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for

XX PT analyzing gene expression in human bone marrow -

XX PS Example 4; SEQ ID NO: 31602; 658bp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC bone marrow. They can be used to measure gene expression in bone marrow

XX CC samples, which may enable the improved diagnosis and treatment of cancers

XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a

XX CC protein encoded by one of the probes of the invention.

XX SQ Sequence 46 AA;

Alignment Scores: 6.14e-11 Length: 46

Pred. No.: 212.00 Matches: 41

Percent Similarity: 91.30% Conservative: 1

Best Local Similarity: 89.13% Mismatches: 4

Query Match: 6.93% Indels: 0
DB: 22 Gaps: 0
US-09-856-061-1 (1-1721) x AAM71296 (1-46)
QY 1239 GATGGTACTTTTGTGTCGAGACTGCTCTACAAAATCCAGGACAGAACCATATGTTTG 1298
|||||:|||||
DB 1 AspGlySerPheIeuValArgAspCySerThrLySerLySGluGluProTyrValIeu 20
QY 1299 GTGGTGTATTATGGGAACAAGGCTACATGTGAAATCCGTTTCCTCGAGACCAATCAA 1358
|||||
DB 21 AlaValPheTyrGluAsnLysValIleArgPheIeuGluArgAsnGln 40
QY 1359 CAGTTGCCCTGGGCACA 1376
|||||
DB 41 GlnPheAlaLeuGlyThr 46

Search completed: April 21, 2003, 12:20:45
Job time : 90.294 secs

TELEX: 910 277299
; INFORMATION FOR CEO ID NO. 1

US-09-856-061-1 (1-1721) x US-08-729-416C-1 (1-474)

```

QY 291 GGATTCGGTATGATGATTCAGAACGTCCTCTCTGTAATAATGATGATCCCAAGC 350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 94 G1yLysSerAsnLeuGlnPheIaGlyMetSerIleSerLeuThrIleSerThrAlaSer 113
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 351 CTCAGCAGTCCCAAGGCGGCTGCGAGCGGCTTCTGGAACCACTCTCCGATCCAGACAG 410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 114 LeuAsnLeuArgThrProAspSerLysGlnIleIle-----AlaAsnHisMetArg 131
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 411 AACTTGGCTGGGTCCCGAGTGGAGAAATGCAACAGTAACAAGCACTAC----- 461
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 SerIleSerPheAlaSerGlyLysPro-----AspThrThrAspArgValAlaTyr 149
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 462 -----GAAATCTCT-----GAGTTCAGCTGCTGAGAGCATGCCATCATAG 503
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 150 ValAlaLysAspProValAsnArgArgAlaCysHisIleLeuGlnCysCysAspGly-- 168
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 504 AAATTTTACCAGCAGCATCCAGATCCAGATCCGATACAGTACAGCTATTTCAG 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 -----LeuAlaGlnAspValIleGlySerIleGlyGlnIleArgIleuArgPheLys 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 564 GATATGATGAGGCTCCCTCTGTGTACTCCCAAGGCTTCTGTCTCCTGAGACAGAA 623
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 G1nTyrLeuGlnCysProThrLysIlePro-----AlaLeuHisAspArgMetGln 203
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 624 ACCGAGATGTGAGATGACAGACTGAGAGAGTGGACAGCAAGCTTCACTTCAGATGTC 683
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 204 SerLeuAspGluProThrLysIleGlu--GluGlnIleLysP----- 215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 684 AGAGCCAGCGCTTTAAAGATTCAAATACCAAAATAAACAAGATCCCTTCCCACT 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 216 -----GlySerAspHisProThrLysAsnSerIleProSerLysMet 229
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 744 CTTGGGCTGTCTATCACTCTCCCAAGAGTACCAACCTTACCCCAAGCAGCAGAG 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 230 ProProProGlyGlyPheLeuAspThrArgLeuLysProArgProHisAlaPro----- 247
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 804 GAGAGCAGTGCATCTCCGCCCAAG----- 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 248 ---AspThrAlaGlnPheIaGlyLysGlnIleThrTyrTyrGlnIleArgHisLeuGly 266
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 831 CCCACCTTTCAGAA---GTCCAGAGGGGGCCAGCAGAGAGTGCAGAAAGACTTCAGT 887
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 267 AspThrPheGlyLysAspTyrGlnIleThrProLeuArgGlnGlySerSetAsp----- 284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 888 AGGGTCTTGGAGCAGAGAAATCTCCACCAAGCAAGCAAGCAAGCAAGCAAGTCTTCCCA 947
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 285 -----IleTyrSerThrProGlyLysLeuHisValAlaProThrGlyLysAlaPro 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 948 TCATCAAAACAAACACACAGAGAGTGCAGCTGCCATCTCCACTCTCTCATATGCA 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 ---ThrTyrValAsnThrGlnIleLeuPro----- 312
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1008 GGAAGACAGATATACAGCCAGAGACATACAGTACAGTACAGTCTGCTGCTAG 1067
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 313 -----GlnAlaTyrProAla----- 317
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1068 AGATGCCAAGCTGCAGCCAGCAGCCCTCGA-----ATGCTGCCCTAT 1112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 ---AlaValSerSerAlaGlySerSerProArgLysAspLeuPheAspMetLysProPhe 336
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1113 GAAACACAAACTCGGAGAAACT----- 1136
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 G1nAspAlaLeuLysAsnGlnProLeuGlyProValLeuSerLysAlaAlaSerValGlu 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1137 -----GACCCCAAAAGCTGATGAGAG-----GATGCTGCGAG 1172
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 357 CysIleSerProValSerProArgAlaProAspAlaLysMetLeuGlnIleuGlnIle 376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1173 AATGATGGTACATTGAGAAATACAGTCCGACAGTGGAGAGTGTAAATGAAGAG 1232
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 377 G1nThrTyrGlnGlyGlnMetSerArgLysGlnAlaGlyLysLeu----- 393
QY 1233 AACAGAGATGATCTTTTGGTCCAGACTCTCTACAAATCCAGACAGACCA--- 1289
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 G1nLysAspGlyAspPheLeuValArg-----LysSerThrAsnProGly 409
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1290 ---TATGTTTGGTGTGTTTATGGCAAGGTCCTACATGTGAAATCCGTTTCTC 1346
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 410 SerPheValLeuThrGlyMetHisAsnGlyGlnAlaLysHisLeuLeuValAspPro 429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1347 GAGAGCAATCAACAGTTTCCCTGGGCAAGACTACGAGGAATGATGTTTATCT 1406
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 430 G1u-----GlyThrIleArgThrLysAspArgValPheAspSer 442
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1407 GTGAGACATCATGATCACTAC-----ACATATTTCCATCTCTCTAATGATGG 1460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 443 IleSerHisLeuIleAsnHisHisLeuGlnSerSerLeuProIleVal----- 458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1461 AAGACAGCTGACGACGAGAAACAGTCACTACCCAGCCACTG 1508
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 459 -----SerAlaGlySerGlnLeuCys---LeuGlnGlnProVal 470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3

```

US-08-729-416C-7
; Sequence 7, Application US/08729416C
; Patent No. 6013767
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
; TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 594 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-729-416C-7

```

Alignment Scores:

```

Pred. No.: 1,84e-07 Length: 594
Score: 165.50 Matches: 106
Percent Similarity: 36.40% Conservative: 60
Best Local Similarity: 23.25% Mismatches: 161
Query Match: 5.41% Gaps: 129
Db: 3

```

US-09-856-061-1 (1-1721) x US-08-729-416C-7 (1-594)

```

QY 291 GCATTCGGTATGATTCAGAAAGCTCTCTGCTGAATAAATAGTCATGCCAACG 350
   ||| :|||:||||| :|||:||||| :|||
Db 214 GYLysSerIAsnLeuGlnPheAlaGlyMetSerIleSerLeuThrIleSerThrAlaSer 233
QY 351 CTACAGCAGTGCACAAAGGGGGGTGTCAGCGGTTCTGAAACCACTTCGGATCAGAAAG 410
   |||:||||| :|||:||||| :|||:||||| :|||
Db 234 LeuAlaSerLeuArgThrProAspSerIleGlnIleIle-----AlaAsnHisMetAlaG 251
QY 411 AACTGGCTGGGCTCCAGCGGTGAGAAATAATGACAGTACACAGCACTAC----- 461
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 252 SerIleSerPheAlaSerIleGlyAlaPro-----AspThrThrAspTyrValAlaTyr 269
QY 462 -----GAAGATCT-----GAGTTCACAGCTGCTGAGAGCAAGCCATCAATG 503
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 270 ValAlaLysAspProValAlaAsnArgAlaCysHisIleLeuGlnCysGlySargly---- 288
QY 504 AAAATTTTACAGCCAGCACTATCCAGAAATGGAATACGCAAGATACAGCTATTTCAG 563
   ||| :||| :||| :||| :||| :|||
Db 289 -----LeuAlaGlnAspValIleGlySerIleGlyGlnIleAlaPheGlnLeuArgPheLys 306
QY 564 GATATGATGAGGCTCCCTCTGTTACTCCCAAGGCTTCTGCTCCATGAGAGACA 623
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 307 GlnTyrLeuGlnCysProThrIlePro-----AlaLeuHisAspArgMetGln 323
QY 624 ACCAGGATGTAGAGATGACACAGCTGAGAGAGTGGACAAAGCTTCAAGATGTC 683
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 324 SerLeuAspGlnProThrPheLys-----GluGlnGlyAsp----- 335
QY 684 ACAAGCCAAAGCTTAAAGATTTCAAAATACACAAATAACAAAGACTCTTGGCAGCT 743
   ||| :||| :||| :||| :|||
Db 336 -----GlySerAspHisProTyrTyrAsnSerIleProSerLysMet 349
QY 744 CCTGGCCTGCTATGATCTCTCCCAAGAGTACAAAGCTTACCCCAAGCAGCAGAG 803
   ||| :||| :||| :||| :|||
Db 350 ProProGlnGlyPheLeuAspThrArgLeuLysProArgProHisAlaPro----- 367
QY 804 GAGACAGTGCATCTGCTGCCAAG----- 830
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 368 ---AspThrAlaGlnPheAlaGlySerGlnThrTyrTyrGlnGlyArgHisLeuGly 386
QY 831 CCCACCTTTCAGAA---GTCAGAGAGGGGCCAGCAGAGAGTGCACAAAGCTTACT 887
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 387 AspThrPheGlnGlyAspThrPheGlnThrProLeuArgGlnGlySerSerAsp----- 404
QY 888 AGGCTCTTGGAGCAGAAAGATCTACACACAGCAGAAAGCCGAATCTTGGCCCA 947
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 405 -----IleTyrSerThrProGlnGlyLysLeuHisValAlaProThrGlyGlnAlaPro 422
QY 948 TCATCAACCAAAACACACAGAGAGTGCACCTGCCATTCGCACAGCTCTTCTACATGCCA 1007
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 423 ---ThrTyrValAsnThrGlnGlnIleProPro----- 432
QY 1008 GGAAGACACTATACAAAGCAGAGACCATACAGGTAGCATGACAGCACTGCTGCTAG 1067
   ||| :||| :||| :||| :|||
Db 433 -----GlnAlaThrProAla----- 437
QY 1068 AGATGCCAAGCTGAGCCAGCCAGCAGCCCTCGA-----ATGCTGCCCTAT 1112
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 438 ---AlaValSerSerAlaGlnSerSerProArgLysAspLeuPheAspMetLysProPhe 456
QY 1113 GAAAACCAAACTCGAGAAACT----- 1136
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 457 GluAspAlaLeuLysAsnGlnProLeuGlyProValLeuSerLysAlaAlaSerValGln 476
QY 1137 -----GACCCCAAAAGCTGATGAGAG-----GATGCTGGCAG 1172
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 477 CysIleSerProValSerProArgAlaProAspAlaLysMetLeuGlnGlnAla 496
QY 1173 AATGAAGGTGATGAGATTAAGTCGCCAGCAGCTGAGAAAGTGGTTAATGAAGAG 1232
   ||||| :||| :||| :||| :||| :|||
Db 497 GluThrTyrPylGlnGlyMetSerArgLysGlnAlaGlnGlyLeuLeu----- 513
QY 1233 AACAAGATGCTACTTTTGGTCCGAGAGCTGCTCAAAATCAAGGAGCAAGCA--- 1289

```

```

Db 514 GYLysAspGlyAspPheLeuValArg-----LysSerThrThrAsnProGly 529
QY 1290 ---TATGTTTGGTGGTGTATATAGGAACAGGCTCTACAAATGTAAATCCGTTCTC 1346
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 530 SerPheValLeuThrGlyMetHisAsnGlyGlnAlaLysHisLeuLeuValAspPro 549
QY 1347 GAGACCAATCAACACTTGGCCCTGGGACAGCAGCTACGAGAAATGAGATGTTGATTC 1406
   ||| :||| :||| :||| :||| :|||
Db 550 Glu-----GlyThrIleArgThrLysAspArgValPheAspSer 562
QY 1407 GGGAAAGCATCATGTAACACTAC-----ACATATTTTCCCATCTCTCTAATGATCGG 1460
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 563 IleSerHisLeuIleAsnHisHisLeuGlnSerLeuProIleVal----- 578
QY 1461 AAAGCAAGCTGCAGCAGCAAGACAGTCTACTACCCAGCCAGCTG 1508
   :|||:||||| :|||:||||| :|||:||||| :|||
Db 579 -----SerAlaGlySerGlnLeuGlySargly-----LeuGlnGlnProVal 590

```

RESULT 4

```

US-08-807-342B-4
; Sequence 4, Application US/08807342B
; Patent No. 6077686
; GENERAL INFORMATION:
; APPLICANT: Der, Channing
; APPLICANT: O'Bryan, John P.
; TITLE OF INVENTION: No. 6077686el SHC Proteins
; NUMBER OF SEQUENCES: 18
; CORRESPONDENCE ADDRESS:
; ADDRESS: Howson and Howson
; CITY: Spring House
; STATE: Pennsylvania
; COUNTRY: USA
; ZIP: 19477
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/807,342B
; FILING DATE: 28-FEB-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/012,516
; FILING DATE: 29-FEB-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Bak, Mary E.
; REGISTRATION NUMBER: 31,215
; REFERENCE/DOCKET NUMBER: MTS2USA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9200
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 550 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-807-342B-4

```

Alignment Scores:

```

Pred. No.: 3 33e-05 Length: 550
Score: 142.00 Matches: 112
Percent Similarity: 34.16% Conservative: 67
Best Local Similarity: 21.37% Mismatches: 203
Query Match: 4.65% Indels: 142
DB: 3 Gaps: 23

```

US-09-856-061-1 (1-1721) x US-08-807-342B-4 (1-550)

Pred. No.:	4.28e-05	Length:	469
Score:	140.50	Matches:	101
Percent Similarity:	34.59%	Conservative:	55
Best Local Similarity:	22.39%	Mismatches:	196
Query Match:	4.60%	Indels:	99
DB:	2	Gaps:	19

OY		1113	GAAACAGCAACTGGGGAACAACSTGC-----CCSACAACCGTGTATGAGAAGAT	1168
Dd		349	GUAABGdlTPRroSerProrotThAtgAdglAlarovaIalArtohtgdlnudn	368
OY		1164	GTCTGCCAAGTAATAAGTGATCATTTGGAGAATTACAGTCSCCAGAGCAATGMAATGTGTA	1223
Dd		369	LLeatUGInclurProTrptUhgIArGmeSeelArgtAladlunArgmetleu	388
OY		1224	AAGAACAGAACAGATAGTAGTATTTTTGGTCSGAGACTGCGCTPACAAATTCAGAGCA	1283
Dd		389	-----AlgAllAsrgLIAsrpheLvalualAgdspservalThr-----AsnpRo	403
OY		1284	GAACGATATGTATTGTCTGTCTTTTATTGGAGAACAGGTCTACAAATGTCAAATCGCTTTC	1343
Dd		404	dGLndltYuallLeuthngLynehlSalaglynlproLyuhlvleuleuleValaspr	423
OY		1344	CITSGAGACATTCACACAGATTGGCTCCCTGGCACAGACTACAGAAAATAGATGTTTGAT	1403
Dd		424	Proluu-----GIyuaIvalArgthrLyuAsprVallLeuheIdnu	436
OY		1404	TCTGTGAAGACATCATTTGAACATACATATTTTCCCATGTCGTATAGATGGGA	1463
Dd		437	SerllleserhnlsleullleAprnhIs-----Leuglnshdeylgn	450
OY		1464	GACAAAGCTGCACCACAGAAAAAGTGSTACSTC	1496
Dd		451	ProIIeValldlaAlaglusegrdluehuIslenu	461
RESULT 8				
		US-08-729-416c-11		
		; Sequence 11 Application US/08729416C		
		: Patent No. 6013767		
		* GENERAL INFORMATION:		
		: APPLICANT: NAKAMURA, TAKESHI		
		: TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE		
		: NUMBER OF SEQUENCES: 22		
		: CORRESPONDENCE ADDRESS:		
		: ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.		
		: STREET: 1100 NEW YORK AVENUE, N.W.		
		: CITY: WASHINGTON		
		: STATE: D.C.		
		: COUNTRY: USA		
		: ZIP: 20005-3918		
		* COMPUTER READABLE FORM:		
		: MEDIUM TYPE: Floppy disk		
		: COMPUTER: IBM PC compatible		
		: OPERATING SYSTEM: PC-DOS/M.S.-DOS		
		: SOFTWARE: Patentin Release #1.0, Version #1.25		
		: CURRENT APPLICATION DATA:		
		: APPLICATION NUMBER: US/08//729,416C		
		: FILING DATE: 11-OCT-1996		
		: CLASSIFICATION: 536		
		: ATTORNEY/AGENT INFORMATION:		
		: NAME: PERRY, GLENN J.		
		: REGISTRATION NUMBER: 28458		
		: REFERENCE/DOCKET NUMBER: 7698/725948		
		: TELECOMMUNICATION INFORMATION:		
		: TELEPHONE: 202-861-3000		
		: TELEFAX: 202-822-0944		
		: TELEX: 67146Z7 CUSH		
		: INFORMATION FOR SEQ ID NO: 11:		
		: LENGTH: 474 amino acids		
		: TYPE: amino acid		
		: TOPOLOGY: linear		
		: MOLECULE TYPE: peptide		
		: US-08-729-416c-11		
Alignment Scores:		8.41e-05	Length:	474
Fred. NO.:		137.50	Matches:	95
score:				

Assignment Scores:		
Pred. No.:	8.41e-05	Length: 474
Score:	137.50	Matches: 95
Percent Similarity:	37.67%	Conservative: 70

Best Local Similarity: 21.69% Mismatches: 180
 Query Match: 4.50% Indels: 93
 DB: 3 Gaps: 23

US-09-856-061-1 (1-1721) x US-08-729-416C-11 (1-474)

QY 291 GGATTCGGTGTATGATGATCCAGAAAGCTCTCTGCTGAAAAAATAGTCATGCCAAGC 350
 ||| :|||:|
 DB 94 GilySerAsnLeuGlnPheAlaGlyMetSerIleSerLeuThrIleSerThrAlaSer 113
 ||| :|||:|
 QY 351 CMCAGAGGCCAAGGGGGGTGTGAGGGGTTCGGAACCACTCCGATACAGAAAG 410
 ||| :|||:|
 DB 114 LeuAsnLeuArgThrProAspSerIleGlnIleIleSer-----AsnHisIleMetArg 131
 ||| :|||:|
 QY 411 AACTGGCTGGGGTCCCGAGTGGAGAAAAATGCACAGTACAGCACTAC----- 461
 :|||:|
 DB 132 SerIleSerPheAlaSerGlyGlyAspPro-----AspThrThrAspArgValAlaTyr 149
 :|||:|
 QY 462 -----GAAGATCCT-----GAGTCCAGCTCTGTAAGGCATGGCCATCAATG 503
 :|||:|
 DB 150 ValAlaLysAspProValAsnArgAlaCysHisIleLeuGlnCysAspGly---- 168
 :|||:|
 QY 504 AAAATTTTACAGCCAGACCTATCCAGAAATGGAATACAGATACAGCTATTTCAG 563
 ||| :|||:|
 DB 169 -----LeuAlaGlnAspValIleGlySerIleGlyGlnAlaPheGlnLeuArgPheLys 186
 ||| :|||:|
 QY 564 GATATGATGAGAGGCTCCCTCTGTTACCTCCCAAGGCTTGTCTCCACTGAGAGACA 623
 :|||:|
 DB 187 GlnTyrLeuGlnCysProSerIlePro-----AlaLeuGlnAspArgMetGln 203
 ||| :|||:|
 QY 624 ACAGAGGATGTGAGATGACACAGCGTGAAGAGTGAACAGCTCACTTCAAGATGTC 683
 :|||:|
 DB 204 SerLeuAspGlnProThrPheGlu---GlnGlnGlyAspGlyPro----- 217
 ||| :|||:|
 QY 684 AGAAGCCACGCTTTAAAGGATTCACAAATACCAAAATAAAGAGACTCCTTGCACCT 743
 :|||:|
 DB 218 -----AspHisProTyrTyrAsnSerValProAsnLysMet 229
 ||| :|||:|
 QY 744 CCTCGGCTGCTATCACTCTCCCAAGAGTACCAACCTTACCCTCCAGACCAACAGAG 803
 ||| :|||:|
 DB 230 ProProProGlyGlyPheLeuAspAlaArgLeuLysAlaArgProHisAlaPro----- 247
 ||| :|||:|
 QY 804 GAGAGCGTCATACCTTCGCTCCAAAGCCACTT----- 839
 :|||:|
 DB 248 AspAlaIleAlaIlePheSerGlyLysGlnGlnThrTyrTyrGlnGlnLysArgIleLeuGlyAsp 267
 :|||:|
 QY 840 -----CCAGAACTCAGAGGGGGCCAGCGAGAGAGTCCAAAGACTTCAGTAGG 890
 :|||:|
 DB 268 AlaPheGlyGlnAspTrpGlnArgAlaProThrArgGlnGlySerLeuAsp----- 284
 ||| :|||:|
 QY 891 GTCTTGAGCAGAGAAGAATCTCACCCACAGCAAAAGCAGAAATCTTCTGCCCATCA 950
 :|||:|
 DB 285 IleTyrSerThrProGlnGlyLysAlaHisMetVal---ProValGlyGlnThrPro-- 302
 :|||:|
 QY 951 TCAAAACAAAACACACAGAGAGTCCACT-----GCCATTGCCAGCTCT 995
 :|||:|
 DB 303 ThrTyrValAsnThrGlnProValProProGlnValTrpProAlaIleThrSerThr 322
 :|||:|
 QY 996 TCCATCATGCCAGGAAGCAGATATACAAAGCCAGAGACCATACAGGTAGCATGCACAC 1055
 :|||:|
 DB 323 GluSerSerProArgLysAspLeuPheAspMetLysProPheGlnAspAlaLeuArgAsn 342
 :|||:|
 QY 1056 TGTCTCTGCTCAGAGATGCCAAGCTGCAGCCAGCCAGCCCTCGAATGCTCCCTATGAA 1115
 :|||:|
 DB 343 GlnProLeuGlyProValIleuSerLysAlaIleSerValGluCysIleSerProVal--- 361
 :|||:|
 QY 1116 AACCAAACTCGGAAACTGCAGCAAGCCCTGATGAGAG-----GATGTC 1166
 :|||:|
 DB 362 -----ThrProArgAlaProAspAlaLysMetLeuGlnGlnLeu 374
 :|||:|
 QY 1167 TGGCAGAAATGATGATTCATTCAGAAATACAGTCCAGCAGAGTGAAGATGTGTAATG 1226
 :|||:|
 DB 375 AsnAlaGlnProTyrPtyGlnGlyGlnMetSerArgLysGlnAlaGlnAlaLeuLeu--- 393
 :|||:|

QY 1227 AAAGAGACAGAGTACTTTTGGTCCGAGACTGCTCTACAAATCCAGGACAGAA 1286
 :|||:|
 DB 394 -----GlnGlnAspGlyAspPheLeuValArg-----LysSerThrThrAsn 407
 :|||:|
 QY 1287 CCA-----TATGTTTGTGGTGGTGTATGGAACAGGTCTACAAATGTAATCCGT 1340
 :|||:|
 DB 408 ProGlySerPheValLeuThrGlyMetHisAsnGlnAlaLysHisIleLeuLeuVal 427
 :|||:|
 QY 1341 TTCCTCGAGCCCAATCAACAGTTGCTCCGGGACAGGACTACGAGAAATGACATGTTT 1400
 :|||:|
 DB 428 AspProGlu-----GlyThrValArgThrLysAspArgValPhe 440
 :|||:|
 QY 1401 GATTCTGTGGAGAACATCATTCATCAACACTAC-----ACATATTTTCCCATCTGCTATA 1454
 :|||:|
 DB 441 AspSerIleSerHisLeuThrThrHisLeuGlnSerSerLeuProIleVal----- 458
 :|||:|
 QY 1455 GATGGGAAAGACAGCTGCACGCGAGAAACAGTCTCACTCCAGCCAGCACTG 1508
 :|||:|
 DB 459 -----SerAlaGlySerGlnLeuLys---LeuArgGlnProVal 470
 :|||:|

RESULT 9
 US-08-729-416C-17
 ; Sequence 17, Application US/08729416C
 ; Patent No. 6013767
 ; GENERAL INFORMATION:
 ; APPLICANT: NAKAMURA, TAKESHI
 ; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
 ; TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREOF
 ; NUMBER OF SEQUENCES: 22
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
 ; STREET: 1100 NEW YORK AVENUE, N.W.
 ; CITY: WASHINGTON
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005-3918
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/729,416C
 ; FILING DATE: 11-OCT-1996
 ; CLASSIFICATION: 536
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: PERRY, GLENN J.
 ; REGISTRATION NUMBER: 28458
 ; REFERENCE/DOCKET NUMBER: 7898/225948
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 202-861-3000
 ; TELEFAX: 202-822-0944
 ; TELEX: 6714627 CUSH
 ; INFORMATION FOR SEQ ID NO: 17:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 594 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-729-416C-17

Alignment Scores:
 Pred. No.: 9 46e-05 Length: 594
 Score: 137.50 Matches: 95
 Percent Similarity: 37.67% Conservative: 70
 Best Local Similarity: 21.69% Mismatches: 180
 Query Match: 4.50% Indels: 93
 DB: 3 Gaps: 23

US-09-856-061-1 (1-1721) x US-08-729-416C-17 (1-594)

QY 291 GGATTCGGTGTATGATTCAGAAAGCTCTCTGCTGAAAAAATAGTCATGCCAAGC 350

```

Db      214 GILYSErsernleuGlnpheAlaGlyMetSerileSerleuThrIleSerThrAlaSer 233
      |||  ::|||::|||  ::|||::|||  |||
QY      351 CTGAGAGTGCACAAAGGGCGGTCTCGAGCGGTCTGTGAAACCACTTCGGATCCAGAAAG 410
      |||:::  ::|||:::  ::|||:::  |||
Db      234 LeuAsnleuArgThrProAspSerLysGlnIleIleSer-----AsnHisIstMetArg 251
      |||:::  ::|||:::  ::|||:::  |||
QY      411 AACCTGGCTGGGTCCAGGTGAGAAATAATGCACAGTAAACACAGTAACTAC----- 461
      ::|||:::  |||::|||  ::|||:::  |||
Db      252 SerileSerpheAlaSerGlyLysAspPro-----AspThrThrAspIleValAlaIleTyr 269
      ::|||:::  |||::|||  ::|||:::  |||
QY      462 -----GAGATCCCT-----GAGTTCACAGCTCTGGAAGCATGGCCATCAATG 503
      ::|||:::  |||::|||  ::|||:::  |||
Db      270 ValAlaLysAspProValAsnArgArgAlaCysHisIleLeuGlnCysCysArgLys--- 288
      ::|||:::  |||::|||  ::|||:::  |||
QY      504 AAAATTTTACAGCCAGACCTATCCAGAAATGGAAATAGCAGATACAGCTATTTCGAG 563
      |||  ::|||  |||  |||
Db      289 -----LeuAlaGlnAspValIleGlySerIleGlyGlnAlaPheGlnLeuArgPheLys 306
      |||  ::|||  |||  |||
QY      564 GATATGATGAGAGGCTCCCTGTGTACTCTCCCAAGGCTTGTGTCTCAGTACAGAGCA 623
      ::|||  ::|||  ::|||  |||
Db      307 GlnTyrLeuGlnCysProSerLysIlePro-----AlaLeuGlnAspArgMetGln 323
      ::|||  ::|||  ::|||  |||
QY      624 ACCAGGATGTAGATGATACAGCTGACAGAAAGTGGACAGACCTTACCTTCAAGATGTC 683
      ::|||  ::|||  ::|||  |||
Db      324 SerLeuAspGlnProTyrThrGlu---GlnGlnGlyAspGlyPro----- 337
      ::|||  ::|||  ::|||  |||
QY      684 AGAAGCCAAAGCTTTAAAGATTCAAATACACAAATAACAGACTCTTGGCACCT 743
      |||  ::|||  |||  |||
Db      338 -----AspHisProTyrTyrAsnSerValProAsnLysMet 349
      ::|||  ::|||  ::|||  |||
QY      744 CCTCGGCTGTATCATCTCTCCCAAGAAATACCAACCTTACCCCAAGCACCAGACAG 803
      |||  ::|||  ::|||  |||
Db      350 PropProTroglyLysPheLeuAspAlaArgLeuLysAlaArgProHisAlaArgPro----- 367
      ::|||  ::|||  ::|||  |||
QY      804 GAGACAGTGCCTACTCTGCTCCCAAGGCCACCTT----- 839
      ::|||  ::|||  ::|||  |||
Db      368 AspAlaIleAlaGlnPheSerGlyLysSerGlnIleThrTyrGlnGlyArgHisLeuGlyAsp 387
      ::|||  ::|||  ::|||  |||
QY      840 -----CCAGAGTCCAGAGGGGGCCAGGAGAGAGTGCACAAAGCTTACGTAGG 890
      ::|||  ::|||  ::|||  |||
Db      388 AlaPheGlyGlnAspTyrGlnArgAlaArgAlaProThrArgGlnGlySerLeuAsp----- 404
      ::|||  ::|||  ::|||  |||
QY      891 GTCCTTGAGAGAGAGAAATCTCCACCAAGCAGACCAAGCAATCTTCTTGGCCATCA 950
      ::|||  ::|||  ::|||  |||
Db      405 IleTyrSerThrProGlnGlyLysAlaHisMetVal---ProValGlyGluThrPro--- 422
      ::|||  ::|||  ::|||  |||
QY      951 TCAAACCAAAACACACAGAGAGTCCACCT-----GCCATTGCCAGCTCT 995
      ::|||  ::|||  ::|||  |||
Db      423 ThrTyrValAsnThrGlnProValProProGlnValTyrProAlaAlaThrSerSerThr 442
      ::|||  ::|||  ::|||  |||
QY      996 TCCATCATGCGCAGAAAGACACAGTATACAGCCAGAGCCATACAGTAGTCAGTACAGC 1055
      |||  ::|||  ::|||  |||
Db      443 GluSerSerProArgLysAspLeuPheAspMetLysProPheGlnAspAlaLeuArgAsn 462
      |||  ::|||  ::|||  |||
QY      1056 TGTCTGCTGCAGAGTGCAGAGTGCAGCCACAGACCCCTGCAATGCTCCCTATGAA 1115
      |||  ::|||  ::|||  |||
Db      463 GlnProLeuGlyProValLeuSerLysAlaHisSerValGlnCysIleSerProVal--- 481
      |||  ::|||  ::|||  |||
QY      1116 AACCAAACTCGGAAAACTGACCCCAACAAAGCTGATGAGAAG-----GATGTC 1166
      |||  ::|||  ::|||  |||
Db      482 -----ThrProArgAlaProAspAlaLysMetLeuGlnGluLeu 494
      |||  ::|||  ::|||  |||
QY      1167 TGGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1226
      |||  ::|||  ::|||  |||
Db      495 AsnAlaGluProTyrTyrGlnGlyGlnMetSerArgLysGlnAlaGlnAlaLeuLeu--- 513
      |||  ::|||  ::|||  |||
QY      1227 AAAGAAACAAAGATGTTTGTGTCGAGACAGTGTCTTACAAATAACAAAGCAGAA 1286
      ::|||  ::|||  ::|||  |||
Db      514 -----GlnGlnAspGlyAspPheLeuValArg-----LysSerThrThrAsn 527
      ::|||  ::|||  ::|||  |||
QY      1287 CCA-----TATGTTTGGTGGTGTGTTTATGGAACAAGTCTACAAATGGAATCCGT 1340
      |||  ::|||  ::|||  |||

```

```

Db      528 ProGlySerPheValLeuThrLysMetHisAsnGlyGlnAlaLysHisLeuLeuVal 547
      |||
QY      1341 TTCCTGAGAGCAATCAACAGTTTGCCCTGGCCACAGCACTACGAGAAATGATGTTT 1400
      |||
Db      548 AspProGlu-----GlyThrValArgThrLysAspArgValPhe 560
      |||
QY      1401 GATTCGTGTGGAAGACATCTTACACACTAC-----ACATATTTTCCATCTGCTAATA 1454
      |||
Db      561 AspSerIleSerHisLeuLeuThrTyrHisLeuGlnSerSerLeuProIleVal----- 578
      |||
QY      1455 GATGGAAAGACAGCTGTCACGCAAGAAACAGTCACTACCCAGCCACTG 1508
      ::|||  ::|||  ::|||  |||
Db      579 -----SerIleGlySerGlnLeuCys---LeuArgGlnProVal 590
      ::|||  ::|||  ::|||  |||

RESULT 10
US-08-434-730-16
; Sequence 16, Application US/08434730
; Patent No. 5637463
;
GENERAL INFORMATION:
;
APPLICANT: Dalton, Stephen
APPLICANT: Kochan, Jarema P
APPLICANT: Osborne, Mark A
TITLE OF INVENTION: METHOD TO DETECT PROTEIN-PROTEIN
TITLE OF INVENTION: INTERACTIONS
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESS: Hoffmann-La Roche Inc
STREET: 340 Kingsland Street
CITY: Nutley
STATE: NJ
COUNTRY: USA
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/434,730
FILING DATE: 04-MAY-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Semionow, Raina
REGISTRATION NUMBER: 39022
REFERENCE/DOCKET NUMBER: 9069
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201)235-4391
TELEFAX: (201)235-2363
INFORMATION FOR SEO ID NO: 16;
SEQUENCE CHARACTERISTICS:
LENGTH: 756 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
US-08-434-730-16

Alignment Scores:
Pred. No.: 0 00021
Score: 134.50 Length: 756
Percent Similarity: 33.178 Matches: 128
Best Local Similarity: 21.448 Conservative: 70
Query Match: 4.408 Mismatches: 204
DB: 1 Indels: 195
Gaps: 29

US-09-856-061-1 (1-1721) x US-08-434-730-16 (1-756)
QY      37 TCGGAAACCAAACTCAACAGCC-----ACATACAGGACACTCTCTGCTGAA 84
      ::|||  ::|||  ::|||  |||
Db      197 AsnSerIleSerHisLeuSerGlyGlyAlaGlyThrValGlyArgAlaLeuAlaAsnAsp 216
      ::|||  ::|||  ::|||  |||
QY      85 GGACTCTGCTGAGGAGAGACATGTCAACTTATCTTACAGAGTGTCCAGATGCA 144
      |||  ::|||  ::|||  |||

```



```

Db 217 GlyThrSerProGlyLysArg----- 223
OY 145 CCGTGGACCCCTTTCCAGAGAGCTACCGCTTCACACTGAGCCCTTGACTAAAGAGA 204
    |||
    ---Trp-----ThHisArgPheGluArgGlyLeuArg 232
Db 224
OY 205 CTGAGCAGGCTGAGTTGAAAGATCCCTCTCTTTTCCAGGTGCCAAGACCATGACAGCC 264
    |||
    LeuSerArg----- 235
Db 233
OY 265 AGGCGCATTAAGACACGAAAGAGATTGCGGTATGAGATTCAGAAAGCTGTC 324
    |||
    ---GlyGlyGlyThrLeuArgAspGlyAlaGlyValIleGlnArgGlyLeu 232
    ---GlyGlyGlyThrLeuArgAspGlyAlaGlyValIleGlnArgGlyLeu 232
OY 325 TCGTGAATAATAGTCATGCGCAAGCTCAGCAGTCCAAAGGCGGTGCGAGCGT-- 382
    |||
    LeuSerPheMetGlyAlaGlyGluAlaAlaProAspProAlaGlyValGlyArgGly 272
    ---TCGGAACCCCTCCGGATCAGACAGAAAGAACTGGGCGGCGTCCAGAGTGGAG 435
    |||
    ---GlyAlaAlaGlyLeuThrSerGlyGly-----GlyGlyGlnProGlnTrpGly 288
    |||
    AAAAATGCAACAGTAAACAGACAGATACGAGATTCAGCTGCTGAAAGCATGCG 495
    |||
    nlysCys-----ArgLeuLeuLeuArgSer-----GluGly----- 298
    |||
    CATCAATGAAATTTTACAGCCAGACCTATCCAGAAATGCGAATTCGACGATACAGCT 555
    |||
    299 -----GluGlyGlyGlySerArg-LeuGlnP 308
OY 556 ATTTCCAGATATGATGAGAGCTCCCTCTGTTACTCCCAAGGCTT----- 605
    |||
    308 hephe-----ValProPolsAlaSerArgProAlaGly 319
    |||
    606 -----GTCTCCACTGAGAGCAAAACAGG 630
    |||
    319 euserIleProCysSerThrIleThrAspValArgThrAlaThrAlaLeuGlnMetProA 339
    |||
    631 ATGTGAGAGTACACAGCTGAGAGAGTGCACAAAGCTTACC-----TTCAAGATG 681
    |||
    339 spArgGluAsnThrPheValValGlyValGlyProSerGluTrpIleLeuGlnThrT 359
    |||
    682 TCAGAGGCAAGCGCTTAAAGATTCAAATACACAAATAACAGATCCCTTTGCCAC 741
    |||
    359 hrAspAlaLeuHisValIleAlaTrp--ValSerAspIleGlnGluCysLeuSerProG 378
    |||
    742 CTCTCGGCTGTATATCACTCTCCCAAGAGTACCAACCTTACCCTCCACACACAG 801
    |||
    378 LyrProCysProAlaIleSer-----ProArgProMetThrLeuPro- 391
    |||
    802 AGGAGAGCATGCTACTCTGCTCCAAAGCCACCTTCCAGAAAGTCCAGAGGGGCCCA 861
    |||
    392 -----LeuAlaProGlyThrSerPhe----- 398
    |||
    862 GGCAGAGAGTGCAAAGACTTACGTAGGGCTTGAGAGCAGAGAAAGATTCACAC 921
    |||
    399 -----LeuThrLysAspAsnThrGluSerLeuGlnLeuProCysLeuAsnHisSerG 416
    |||
    922 AGACAAAGCCAGAAATCTTCT-----TGCCCATCATCAAAACAAACACAGAGAGTGC 975
    |||
    416 luserLeuProSerGlnAspLeuLeuGlnGlyProSerGluSerAsnAspArgLeuSerG 436
    |||
    976 CACCTGCATTCGCC-----AGCTTTCTTCTACTGCCAGAGAAAGCACA 1017
    |||
    436 lnglAlaIleGlyGlyLeuSerAspArgProSerAlaSerPheSerProSerAlaAs 456
    |||
    1018 GATATCAAGCCAGAGACATATAGTAGATGACAG-----C 1053
    |||
    456 etIle-----AlaAlaSerHisPheAspSerMetGlnLeuLeuProGluLeuProProA 475
    |||
    1054 ACTGTCCTGCTCAGAGATGCGCAAGCTGACGCGCAGCAGACCCCTCGAATGCGCCTATG 1113
    |||
    475 rgIleProIleGlnGlnGlyProProAlaGlyThrValHisProLeuSerThrProTyrP 495

```

```

OY 1114 AAAACAAACTCGGAGGAAGAACTGACCCCAAGAGCTGATGAGAAGATGCTGCAGA 1173
    |||
    495 roProLeuAspThr-----ProGluAlaAlaThrGlySerPheLeuPheGlnG 511
    |||
    1174 ATGAA-----TCGTACATTTGAG 1191
    |||
    511 lylGluAlaGlnGlyGlyGlnGlyAspGlnProLeuSerGlyTrpTrpProPheHisGlyM 531
    |||
    1192 AATACAGTCCGACGAGCAGTGAAGATGCTGTAATGAAA-----GAGAACAGATGTA 1245
    |||
    531 etLeuSerArgLeuLysAlaAlaGlnLeuValLeuGlnGlyThrSerHisGlyVal 551
    |||
    1246 CTTTTCGTCGAGAGCTGCTTACAAATTCAGAGCAGACACATGTTTGGTGTGTGT 1305
    |||
    551 alPheLeuValArgGlnSerGluThrArg--ArgGlyGln--TyrValIleThrPheA 569
    |||
    1306 TTTATGGGAACAGTCTCATATGTGAAATCCGTTTCTCGAGACCAATCAAGTTTG 1365
    |||
    569 snPheGlnGlyLysAlaLysHisLeuArgLeuSerLeuAsnGlnGlyGln----- 586
    |||
    1366 CCTGGGACACAGACTACAGAGAAATGAGATG--TTGATTCGTGGAACATCATG 1422
    |||
    587 -----CysArgValGlnHisLeuThrPheGlnSerIlePheAspMetLeuG 602
    |||
    1423 AACACTACACATATTTCCATCTGCTAATAGATGG----- 1460
    |||
    602 luhisPheArgValHisProIleProLeuGlnSerIleGlySerSerAspValIleLeuV 622
    |||
    1461 -----AAGACAGGCTGCACGACGAGAAACAGTCTTACC 1494
    |||
    622 alserTyrValProSerGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnG 642
    |||
    1495 TCACCCAGCCACTGCTCTGCGCCAGGCTCTTCTCACTAGTACTCAGCAGGACTTC 1554
    |||
    642 roserGluProProPro-----TrpHisAspProProHisProGlyAlaG 657
    |||
    1555 ATGAGTAAAGCCGACGATATCCCGCATCGAGTGGGCGCC 1601
    |||
    657 lnglAlaSerGlyAlaProGluValAlaAlaAlaThrAlaAlaAla 672

```

RESULT 11
 US-08-980-080-2
 ; Sequence 2, Application us/08980080
 ; Patient No. 6312941
 ; GENERAL INFORMATION:
 ; APPLICANT: CARTER-SU, CHRISTIN
 ; APPLICANT: RUI, LIANG-YOU
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR IDENTIFYING
 ; TITLE OF INVENTION: SIGNALING PATHWAY ACONISTS AND ANTAGONISTS
 ; NUMBER OF SEQUENCES: 6
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: MEDLEN & CARROLL, LLP
 ; STREET: 220 MONTGOMERY STREET, SUITE 2200
 ; CITY: SAN FRANCISCO
 ; STATE: CALIFORNIA
 ; COUNTRY: UNITED STATES OF AMERICA
 ; ZIP: 94104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/980,080
 ; FILING DATE: 26-NOV-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: CARROLL, PETER G.
 ; REGISTRATION NUMBER: 32,837
 ; REFERENCE/DOCKET NUMBER: UM-03036
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 705-8410

```

; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 670 amino acids
;   TYPE: amino acid
;   STRANDEDNESS: not relevant
;   TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-09-856-061-1 (1-1721) x US-08-980-080-2 (1-670)

Alignment Scores:
  Pred. No.: 0.00022
  Score: 134.00
  Percent Similarity: 33.71%
  Best Local Similarity: 22.47%
  Query Match: 4.38%
  Gaps: 27

  Length: 670
  Matches: 120
  Conservative: 60
  Mismatches: 180
  Indels: 174
  Gaps: 27

US-09-856-061-1 (1-1721) x US-08-980-080-2 (1-670)
  37 TCGAANAACCAAACTCAACAGC-----ACATACAGGCACTCTCTGTGA 84
  197 AanserAmsrserAmsrserelgylaglythrvallgylvalleuAlaasnsp 216
  85 GCACCTCTGTGAGGGAGAAACATGTCACTTATCTTACAGAGTGTCTCAGATGCGA 144
  217 GlythrsrProgllygluArg----- 223
  145 CCGTGGACCCCTTCCAGAGCTAGCCGTCTACACACTGAGCCCTTGACTAAAGAGA 204
  224 ---Trp-----ThlHsArghPhegluArgleuArg 232
  205 CTGACAGGCTGAGTTGAAGATCCCTCTTTTCCAGGTGCCAAGACATGACAGCC 264
  233 LeuserArg----- 235
  265 AGGCAATAAAGAGCAAGAAAGAGTTCCGTGATCTGAGATCCAGAACCTCTCTC 324
  236 -----GlyGlyGlyThrleuArgAspGlyAlaGlyValIleGlAlaGluLeu 252
  325 TCGTGAANAATAGTATGCGCCACCCCTCAGCAGTGCCTGAGGCGGTGAGCGGT-- 382
  253 LeuserPheMetGlyAlaGluLeuAlaAlaProAspProAlaGly-ValIleArgGlyG1 272
  383 -----TCTGGAACCACTTCGCGATCACAGAGAACTGCTGCGGTGCCAGGTGAG 435
  272 yGlyAlaIleGlyLeuThrSerGlyGly-----GlyGlyGlnProGlnTrpG1 288
  436 AAAAATGCACAGTAAACAGACTACGAAGATCTGAGTTCAGCTGCTGAAGCATGCG 495
  288 nlyecys-----ArgleuLeuLeuArgSer-----GluGly----- 298
  496 CATCAATGAATTTTACCAGCCAGACCTATCCAGGAATCGGAATTCAGCATACAGCT 555
  299 -----GluGlyGlyGlyGlySerArg-LeuGluP 308
  556 ATTTCAGGATATGATGAGGCTCCCTTGTACCTCCCAAGGCTTCT----- 605
  308 hephe-----ValProPolysAlaSerArgProArgL 319
  606 -----GTCTCCACTGAGAGCAACACAGG 630
  319 euserIleProCysSerThrIleHrsPValArgThrlAlaThrAlaLeuGlnMetProA 339
  631 ATGAGAGATGACACACTGGAAGAAAGTGCACAGCTTACC-----TTCAAGATG 681
  339 sPArgGluAsnThrPheValValIlyValIlyGluGlyProSerGluTrpIleleuGlnThrT 359
  682 TCAGAACCAACCCCTTAAAGAGATTCAATACAAATAAACAAACACACCTTTGGCAC 741
  359 htrAspAlaLeuHnIleValIleAlaTrp---ValSerAspIleGlnGluCysLeuSerProG 378
  742 CTCCTGGCCCTGTACTCTCTCCCAAGAAAGTACCAACCTTACCCCGACACACAG 801

```

```

  378 lYProCysProAlaIleSer-----ProArgProMetThrLeuPro- 391
  802 AGGAGAGCAGTGCATCTTCGCTCCAAAGCCACCTTTCAGAAATCCAGAGGGGCCCA 861
  392 -----LeuAlaProGlyThrSerPhe----- 398
  862 GGCAGAGGAGTGCAAAAGCTTCACTAGGCTCTTGAGCAGACAGAAATCTCACACC 921
  399 -----LeuThrLysAspAsnThrGlnSerLeuGlnLeuProCysLeuAsnHnIleSerG 416
  922 ACACAAGCCAGAAATCTTCT-----TGCCCATCATCAAAACCAACACAGAGACTG 975
  416 luserLeuProSerGlnAspLeuLeuGlyProSerGlnuserAsnAspArgLeuSerG 436
  976 CACCTGCCATTGSC-----AGCTCTCTCATGTCAGGAAAGCA 1017
  436 lnglyAlaIleGlyGlyLeuSerAspArgProSerAlaSerPheSerProSerAlaAs 456
  1018 GTATCAAGCCAGACCATACAGTACATGACAG-----C 1053
  456 erIle---AlaAlaSerHnIlePheAspSerMetGluLeuLeuProProGluLeuProPro 475
  1054 ACTGCTCTGTGAGATGCCAGAGCTGCAGCCAGCCACAGCCCTCGAATGCTGCCCTATG 1113
  475 rgIleProIleGluGluGlyProProAlaGlyThrValHnIleProLeuSerThrProTy 495
  1114 AAACACAACCTCGGAGAAACCTGACCCCAAGCCCTGATGAGAAGATGTCTGCAG 1173
  495 roProLeuAspThr-----ProGluAlaIleThrGlySerPheLeuHnGln 511
  1174 ATGA-----TGATACATTGGAG 1191
  511 lylGluAlaGlyGlyGlyGlyGlyAspGlnProLeuSerGlyTyrrProThrPheHnIle 531
  1192 AATACAGTCCGACGAGCAGTGCAGATGTGTATGAAA-----GAGACCAAGATGTA 1245
  531 etLeuserArgLeuLysAlaIleGlnLeuValIleGluGlyGlyThrSerSerHnIleGly 551
  1246 CTTTTCCTGCTCCGACACTGCTCTCAAAATCCAAAGCAGAAACCATATGTTGGTGTG 1305
  551 alPheLeuValArgGlnSerGluThrArg--ArgGlyGlu--TyrrValLeuThrPheA 569
  1306 TTTATGGGAACAGCTATCAATGTGAATAATCCGTTTCTCGAGAGCAATCAACAGTTTG 1365
  569 snPheGlnGlyLysAlaIleLysHnIleuArgLeuSerLeuAsnGluGlyGln----- 586
  1366 CCTGGGCACAGGACTACAGAGAAATGAGATG--TTGATTCGTGGAAGACATCATTG 1422
  587 -----CysArgValGlnHnIleuThrPheGlnSerIlePheAspMetLeuG 602
  1423 AACACTACATATTTTCCCATCTGCTATATACATGGG 1460
  602 lnhIlePheArgValHnIleProIleProLeuGlnuserGly 614

RESULT 12
US-08-353-550-8
; Sequence 8, Application us/08353550
; Patent No. 5744313
; GENERAL INFORMATION:
; APPLICANT: Williams, Lewis T.
; APPLICANT: Cavanaugh, William M.
; TITLE OF INVENTION: No. 5744313el Protein Domain Which Binds
; TITLE OF INVENTION: Tyrosine Phosphorylated Proteins
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend Kourile and Crew
; STREET: One Market Plaza, Steuart Tower, Suite 2000
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94105
; COMPUTER READABLE FORM:

```

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/353,550
 FILING DATE: 09-DEC-1994
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Dow, Karen B.
 REGISTRATION NUMBER: 29,684
 REFERENCE/DOCKET NUMBER: 02307k-057300
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: 415-326-2400
 TELEFAX: 415-326-2422
 INFORMATION FOR SEQ ID NO: 8:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 474 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-353-550-8

Alignment Scores:
 Pred. No.: 0.000205 Length: 474
 Score: 133.50 Matches: 119
 Percent Similarity: 35.11% Conservative: 72
 Best Local Similarity: 21.88% Mismatches: 231
 Query Match: 4.37% Indels: 122
 DB: 1 Gaps: 26

US-09-856-061-1 (1-1721) x US-08-353-550-8 (1-474)

QY 17 CAGGTCTGTGGCTGCTTCTCGAAACCAAACTCAA-----CAG 58
 Db 3 LysLeuSerGlyGlyGlyArgGlyArgValGluGlyGlnLeuGlyGlyGlu 22
 QY 59 GCACATACAGGCGC---TCCTGCTGAAGAGCTGCTGAGGGAGAACATGTCAC 115
 Db 23 GluTrpThrArgHisGlySerPheValAsnLysProThrArgLysTrpLeuHisProAsn 42
 QY 116 -----TCATCTTACAGAGTGTCTCCAGATCGACCGCTGA 151
 Db 43 AspLysValMetGlyProGlyValSerTyrLeuValArgTyrMetGlyCys----- 59
 QY 152 CCCCCTTCCAGAGCTAGCCGCTCTACACTGAGCCCTTGACTAAAG---AAGACTGA 208
 Db 60 -----ValGluValLeuGlnSerMetArg-AlaLeuAspPheAsnThrArgThrGlu 76
 QY 209 GCAGGCTGAGTGAAGATCCCTCTTTGCGAG---GTGCCAAGACATGACACGCA 265
 Db 76 nValThrArgGluAlaIleSerLeuValCysGluAlaValPro----- 90
 QY 266 GGGCAATMAAGACCAAGAA----- 290
 Db 91 -GlyAlaIleGlyAlaThrArgArgArgLysProCysSerArgProLeuSerSerIleLe 110
 QY 291 -GGATTGGGATGTGATTCAGAACGTCCTCTGCTGAAAAATAGTCATGCGCAAG 349
 Db 110 uGlyArgSerAsnLeuLysPheAlaGlyMetProIleThrLeuThrValSerThrSerSe 130
 QY 350 CCTGAGAGTGCCTGAAAGGGGCTGTCAGCGGTTCTGAAACCACTTCCGATCCACAG 409
 Db 130 rLeuAsnLeuMetAlaIleAlaIleAspCysLysGlnIleLe-----AlaAsnHisMetGlu 148
 QY 410 GAACCTTGCTGGGTCCAGAGTGAAGAAAAATGCAACAGTAAACAGTAACTGCAAGATCC 469
 Db 148 nSerIleSerPheAlaSerGlyGly-----AspPr 158
 QY 470 TGAGTTCAGGCTGCTGAAGAGCGCATCATGAAATTTTACAGCAGACCTATCA 529
 Db 158 oAspThrAlaGluTyrValAlaIleTyrValAlaIleLysAspProValAsnGlnArgAlaCysHis 178

QY 530 GGAATCGAATATACGAGATACAGCTATTTCAGGATATGAGAGCTCCCTTCTGT 589
 Db 178 s-----IleLeuGlnCysProGluGlyLe 186
 QY 590 ACCGCCAAGGCTTGTCTCTCACTGAGAGACAAACAGGATGTGAGATGACACAG-- 647
 Db 186 uAlaGlnAspVal---IleSerThrIleGlyGlnAlaPheGlnLeuLeuArgPheLysGlnTyr 205
 QY 648 -CTGAGAGAGTGAACAGACCTTCAAGATTCAGAGATTCGAAAGCCCAACGCTTAAAGATT 706
 Db 205 rLeuArgAsnProProLysLeuValThrProHisAspArgMetAlaGlyPheAspGlySe 225
 QY 707 CAATATACAAATAATAAACAGACTCTTGGCCACCTCTCGGCTGCTATCACCTCC 766
 Db 225 rAlaTrpAspGlnGluGlnGluGluPro---ProAspHisGlnTyrTyrAsnAspPhePr 244
 QY 767 CAAGAATACCAACCTTACCCCAAGCACCACCA-----GAGAGAGCAGTGCATA 817
 Db 244 oGlyLysGlnProProLeuGlyGlyValAlaPheMetArgLeuArgGluGlyAlaIleArg 264
 QY 818 CTTCCGCTCAAGCCCACTTCCAGAACTCCAGAGGGGGCCCGCAGAGAGAGTCAAA 877
 Db 264 oGlyAlaAlaArgProThrAlaProAsnAlaGlnThrProSerHisLeuGlyAlaThrLe 284
 QY 878 AGACTTCAGTAGGCTCTTGAGAGAGAGAAATCTCACCAAGCAACCAAGCAAGATC 937
 Db 284 uProValIleGlnProValIleGlyLysProGluValArgLysGlnMetProProPro 304
 QY 938 TTCTGGCCCATCATCAACCAAAACACACAGAGATCCACCTGCCATTGCCAGCTTTC 997
 Db 304 oProCysProAlaGlyArgGluu-----LeuPheAspAspProSe 317
 QY 998 CTGATCCAGAGAAACAGATATACAGCCAGAGACCATATACAGATGAGCATGAGACTG 1057
 Db 317 rTyrValAsnValGlnAsnLeuAspLysAlaArgGlnAlaValGlyGly----- 333
 QY 1058 TCCTGCTCAGAGATGCAAGCTGTCACCGACCGACCCCTCGA-----ATGCT 1105
 Db 334 -AlaGlyProProAsnProAlaIleAsnGlnSerAlaProArgAspLeuPheAspMetLys 353
 QY 1106 GCCCTTATGAACACAAACTCGAGAAACCTGACCCCAACAAAGCTGATGAGAGATGT 1165
 Db 353 sProPheGlnAspAlaLeuArgValProProProGlnSerValSerMetAlaIleLys 373
 QY 1166 CTGCGAATGAA---TGATACATTGAGAGATACAGTCCGAGAGCTGGAAGATGTGT 1222
 Db 373 nLeuArgGlyGlnProTrpPheHisGlyLysLeuSerArgArgGluAlaIleLeuLe 393
 QY 1223 AATGAAGAGACAGAGATGCTTTTGTGCGAGACTGCTTACAAATCCAAAGC 1282
 Db 393 u-----GlnLeuAsnGlnLysPheLeuValArgGlnSerThrThr-----ThrPr 408
 QY 1283 AGAACCATATGTTTGTGTGTTTATATGGAACAAGTGTACAAATGTAATCCGTTT 1342
 Db 408 oGlyGlnTyrValLeuThr-----Gln 415
 QY 1343 CTTGAGAGCAATCAACAGTTTCCCTG-----GGCAGAGACTAGC 1384
 Db 415 yLeuGlnSerGlyGlnProLysHisLeuLeuValAspProGluGlyValAlaArgThr 435
 QY 1385 AGCAATAGAGATGTTGATTTCTGTGAGACATCATTTGAACACTAC-----ACATATT 1438
 Db 435 rLysAspHisArgPheGlnSerValSerHisLeuSerTyrHisMetAspAsnHisLe 455
 QY 1439 TCCCATCTGCTATATAGATGAGAAAGAGCGTCCAGCAGAAACAGTCTACCTCAC 1498
 Db 455 uProIleIle-----SerAlaGlySerGlnLeuLys-----LeuGln 467
 QY 1499 CCAGCCACTG 1508
 Db 467 nGlnProVal 470

RESULT 13


```

Db 408 OGlyGlnTyrValLeuThr-----G1 415
QY 1343 CCTGAGAGCAATCAACAGTTTGCCCTG-----GGCAGAGACTACG 1384
    |||:||||| |||
Db 415 yLeuGlnSerGlyGlnProLysHisLeuLeuValAspProGluGlyValArgTh 435
QY 1385 AGCAATAGAGTGTTCATTCTGTGGAAGACATCATTTGAACACTAC-----ACATATT 1438
    :|||:||||| |||:|||||
Db 435 rLysAspHisArgPheGlnSerValSerHisLeuIleSerTyrHisMetAspAsnHisLe 455
QY 1439 TCCCATCTGTGTATATAGATGGAAAGACAGGCTGCACGCGAAGAAAGTGCCTACCTCAC 1498
    |||:|||||: |||:|||||
Db 455 uProIleIle-----SerIleGlySerGluLeuGlyS--LeuG1 467
QY 1499 CCAGCCACTG 1508
    |||:|||||:
Db 467 nGlnProVal 470

RESULT 14
US-08-426-509A-4
; Sequence 4, Application US/08426509A
; Patent No. 6326469
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Gishizsky, Mikhail
; APPLICANT: Sures, Imman G.
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
; TITLE OF INVENTION: TYROSINE KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York,
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,509A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/232,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-0074-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 675 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: NO. 6326469e
; US-08-426-509A-4

Alignment Scores:
Pred. No.: 0.000276 Length: 675
Score: 133.00 Matches: 62
Percent Similarity: 39.92% Conservative: 41
Best Local Similarity: 24.03% Mismatches: 107
Query Match: 4.35% Indels: 48
DB: 4 Gaps: 11

```

```

US-09-856-061-1 (1-1721) x US-08-426-509A-4 (1-675)
QY 750 CCTGCTATTCACCTCTCCCAAGAAATACCAACCTTACCCCCAGACACAGAGAGAGC 809
    |||:||||| |||:|||||
Db 141 ProGlyCysThrLeuTyrGlnAlaTyrAlaAsnLeuHisThrAlaValAsnGluGlnLys 160
QY 810 AGTCGATACCTTCGCTCCAAAGCCACTTCCAGAA--GTCCAGAGGGGGCCAGG--- 863
    |||:||||| |||:|||||
Db 161 HisArgVal-----ProThrPheProAspArgValLeuLysIleProAlaGala 176
QY 864 -----CAGAGAGTGCMAAAGACTTCAGTAGGGTCTCTTGAGACA----- 902
    :|||:||||| |||:|||||
Db 177 ValProValLeuLysMetAspAlaProSerSerSerThrThrLeuAlaGlnTyrAspAsn 196
QY 903 GAAGAGAAATCTCACCAACAGACAGACCAAGCAAGATCTTCTGCGCATATCAAAACCAAAAC 962
    |||:||||| |||:|||||
Db 197 GlnSerLysLysAsnTyrGlySerGlnProProSerSerSerThrSerLeuAlaGlnTyr 216
QY 963 ACACAGAAAGTCCACCTGCCATTCGCCAGCTCT-----TCTACATGCCA 1007
    |||:||||| |||:|||||
Db 217 AspSerAsnSerLysLysIleTyrGlySerGlnProAsnPheAsnMetGlnTyrIlePro 236
QY 1008 GGA-----AAGCAGATATACAGCCAGA 1031
    |||:||||| |||:|||||
Db 237 ArgGlnAspPheProAspTrpTrpGlnValArgLysLeuLysSerSerSerSerGlu 256
QY 1032 GACCATACAGGTAGCATGACAGCAGCACTGCTCTCTCAGAGATCCCAAGTCGACGACGAC 1091
    |||:||||| |||:|||||
Db 257 AspValAlaSerSerAsnGlnLys-----GluArgAsnValAsnHis 270
QY 1092 AGCCCTGGAATGTCGCCCTATGAAACACAAATCTCGAGAAACCTGACCCCAAAAGCCT 1151
    |||:||||| |||:|||||
Db 271 ThrThrSerLysIleSerTrpGlnPheProGlnSerSerSerGlu----- 286
QY 1152 GATGGAAGAGATGTCGCGACAAATGATGATGAGATTCGAGAAATACAGTCCGACGAGG 1211
    |||:||||| |||:|||||
Db 287 GlnGlnGlnAsnLeuAspAspTyrAspTrpPheAlaGlnAsnIleSerArgSerGlnSer 306
QY 1212 GAAGATCTGTATATGAAGAAGACAGACAGATGTAAGTCTTTGGTCCGAGCTGCTTACA 1271
    |||:||||| |||:|||||
Db 307 GlnGlnLeuLeuArgGlnGlyGlnGlyGlnGlyAlaPheMetValArgAsn----- 323
QY 1272 AAATCCAGAGCAGAACCATATGTTGGTGGTGT-----TATGGACAAAGCTC 1322
    |||:||||| |||:|||||
Db 324 SerSerGlnValGlyMetTyrThrValSerLeuPheSerLysAlaValAspLysLys 343
QY 1323 TACAATGTGAATAATCCGTTTCTCGAGAGCAATGACAGTTTGCCCTGGCAGAGACTA 1382
    |||:||||| |||:|||||
Db 344 GlyThrValLysHisTyrHisValHisThrAsnAlaGlnAsnLysLeu-----TyrLeu 361
QY 1383 CGAGAAATGAGATGTTGATTCGTGGAAGACATCATTTGAACACTACATAT 1436
    |||:||||| |||:|||||
Db 362 AlaGlnAsnTyrCysPheAspSerIleProLysLeuIleHisTyrHisGlnHis 379

RESULT 15
PCT-US95-05008-4
; Sequence 4, Application PC/TUS9505008
; GENERAL INFORMATION:
; APPLICANT: Sugen, Inc.
; APPLICANT: 515 Galveston Drive
; APPLICANT: Redwood City, California 94063-4720
; APPLICANT: United States of America
; APPLICANT: Hoffgarten E.V.
; APPLICANT: Hoffgarten Str. 2
; APPLICANT: Munchen 80539
; APPLICANT: Germany
; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
; TITLE OF INVENTION: Kinases
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas

```

```

CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNITE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 675 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-05008-4

```

```

Alignment Scores:
Pred. No.: 0.000276 Length: 675
Score: 133.00 Matches: 62
Percent Similarity: 39.928 Conservative: 41
Best Local Similarity: 24.038 Mismatches: 107
Query Match: 4.358 Indels: 48
DB: 5 Gaps: 11

```

US-09-856-061-1 (1-1721) x PCT-US95-05008-4 (1-675)

```

QY 750 CCGTATACCTCTCCCAAGAGTACCAACCTTACCCCAAGACACCAAGAGAGAC 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 141 ProGlyCystrLeuTrpGluIaTyraIaAsnLeuHisThrAlaValasngluLys 160
QY 810 AGTGCACTGCGTCCAAAGCCACCTTCCAGAA---GTCCAGAGGGGCCAGG--- 863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 161 HisArgVal-----ProThrPheProAspArgValLeuLysIleProArgAla 176
QY 864 -----CAGAGAGTGCACAAAGACTTCAGTAGGGTCTTGAGACA----- 902
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 ValProValLeuLysMetAspAlaProSerSerSerThrThrLeuAlaGlnTyrAspAsn 196
QY 903 GAAGAAGATCTCACCAACAGCAAGCAAGATCTTCTGCCATCATCAACCAAAAC 962
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 GluSerLysLysAsnTyrLeuSerGlnProProSerSerSerThrSerLeuAlaGlnTyr 216
QY 963 ACACAGAGAGCTCCACCTGCATTCAGCACTCT-----TCCATCATGGCCA 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 217 AspSerAsnSerLysLysIleTyrGlnProAsnPheAsnMetGlnTyrIlePro 236
QY 1008 GGA-----AACACAGTATACAAAGCCACA 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 237 ArgGluAspPheProAspTrpTrpGlnValArgLysLeuLysSerSerSerSerGlu 256
QY 1032 GACCATACAGTAGTACAGACACTGCTGCTCAGAGATGCCAAGTGCAGCCAC 1091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 257 AspValAlaSerSerAsnGlnLys-----GluArgAsnValAsnHis 270
QY 1092 AGCCCTCGAATGCTGCCCTATGAAACACAAACTCGAGAAACCTGACCCACAAAGCCT 1151

```

```

DB 271 ThrThrSerLysIleSerTrpGluPheProGluSerSerSerSerGlu----- 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1152 GATGAGAAGATGTCTGCGACAAATGAATGATGTACATTGAGCAATATACAGTCCAGGAGTGTG 1211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 287 GluGluGluAsnLeuAspAspTyrAspTrpPheAlaGlnLysIleSerArgSerGlnSer 306
QY 1212 GAAGATGTGTTAATGAAGAAGAACAGAGATGACTTTTGTGTCGAGACTGCTTACA 1271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 307 GluGlnLeuLeuArgGlnLysGlyLysGlyLysGlyAlaPheMetValArgAsn----- 323
QY 1272 AAATCCAGGACAGAACCATATGTGTTGGTGTGCTGTT-----TATGGAAACAAAGTTC 1322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 324 SerSerGlnValGlyMetTyrThrValSerLeuPheSerLysAlaValAsnAspLysLys 343
QY 1323 TACAAATGTGAATAATCCGTTTCTCGAGAGCAATACACAGTTTCCCTGGGACAGAGACTA 1382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 344 GlyThrValLysHisTyrHisValHisThrAsnAlaGlnAsnLysLeu-----TyrLeu 361
QY 1383 CGAGAAATGAGATGTTTGTATCTGTGGAAGACATCATTCAGACATAT 1436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 AlaGluAsnTyrCysPheAspSerIleProLysLeuLeuHisTyrHisGlnHis 379

```

Search completed: April 21, 2003, 12:32:57
 Job time : 39.2293 secs

GenCore version 5.1.4-P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 12:20:58 ; Search time 43.7798 Seconds

(without alignments)
5944.816 Million cell updates/sec

Title: US-09-856-061-1

Perfect score: 3057

Sequence: 1 acagagcccaactgcacag.....taaaaaaaaaaaaaa 1721

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 288829 segs, 75613885 residues

Total number of hits satisfying chosen parameters: 577658

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame-n2p.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool/US09856061/runat_21042003_113021_3809/app_query.fasta.1.3150
-DB=Published.Applications_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=biosum62
-TRANS-human40.cdt -LIST=45 -LOCALIGN=200 -THR.SCORE=pct -THR.MAX=100
-THR.MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pro -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09856061 -CCGN_1.1_34 -runat_21042003_113021_3809
-NCPU=6 -ICPU=3 -NO_XLPHY -NO_MMAR -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGCLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications_AA:*

1: /cgn2.6/prodata/2/pubppa/US08_NEW_PUB.pep:*
2: /cgn2.6/prodata/2/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2.6/prodata/2/pubppa/US06_NEW_PUB.pep:*
4: /cgn2.6/prodata/2/pubppa/US06_PUBCOMB.pep:*
5: /cgn2.6/prodata/2/pubppa/US07_NEW_PUB.pep:*
6: /cgn2.6/prodata/2/pubppa/US07_PUBCOMB.pep:*
7: /cgn2.6/prodata/2/pubppa/PCTUS_PUBCOMB.pep:*
8: /cgn2.6/prodata/2/pubppa/US08_PUBCOMB.pep:*
9: /cgn2.6/prodata/2/pubppa/US09_NEW_PUB.pep:*
10: /cgn2.6/prodata/2/pubppa/US09_PUBCOMB.pep:*
11: /cgn2.6/prodata/2/pubppa/US10_NEW_PUB.pep:*
12: /cgn2.6/prodata/2/pubppa/US10_PUBCOMB.pep:*
13: /cgn2.6/prodata/2/pubppa/US60_NEW_PUB.pep:*
14: /cgn2.6/prodata/2/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	1370.5	44.8	44.3	9	US-09-966-955A-2
2	1316.5	43.1	428	9	US-09-966-955A-4
3	1109.5	36.3	353	9	US-09-966-955A-6
4	212	6.9	46	10	US-09-864-761-38652

5	136.5	4.5	370	9	US-09-764-668-764	Sequence 764, App
6	133	4.4	675	9	US-10-186-399-3	Sequence 3, Appl1
7	133	4.4	675	9	US-09-977-260-4	Sequence 4, Appl1
8	133	4.4	675	9	US-09-977-261-4	Sequence 4, Appl1
9	133	4.4	675	10	US-09-977-269-4	Sequence 1, Appl1
10	133	4.4	688	9	US-10-081-980B-1	Sequence 1, Appl1
11	133	4.4	724	9	US-10-081-980B-3	Sequence 3, Appl1
12	133	4.4	724	10	US-09-962-929-4	Sequence 4, Appl1
13	130	4.3	724	10	US-09-962-929-2	Sequence 2, Appl1
14	128.5	4.2	511	10	US-09-864-761-34590	Sequence 34590, A
15	128	4.2	511	10	US-10-186-399-2	Sequence 2, Appl1
16	128	4.2	1317	10	US-09-963-896-7	Sequence 7, Appl1
17	125	4.1	524	9	US-09-764-868-761	Sequence 761, App
18	124.5	4.1	593	10	US-09-920-021A-3	Sequence 22, Appl1
19	124.5	4.1	597	9	US-10-038-010-22	Sequence 10, Appl1
20	123.5	4.0	1056	9	US-10-161-510-10	Sequence 38, Appl1
21	123.5	4.0	1736	10	US-09-919-97-98	Sequence 9, Appl1
22	122.5	4.0	1163	9	US-09-932-257A-3	Sequence 3, Appl1
23	121	4.0	1325	10	US-09-864-761-35612	Sequence 35612, A
24	121	4.0	2442	12	US-10-109-886-10	Sequence 10, Appl1
25	119	3.9	659	9	US-09-977-260-8	Sequence 8, Appl1
26	119	3.9	659	9	US-10-045-202-2	Sequence 2, Appl1
27	119	3.9	659	9	US-09-977-261-8	Sequence 8, Appl1
28	119	3.9	659	10	US-09-977-269-8	Sequence 8, Appl1
29	118.5	3.9	526	10	US-09-801-368-362	Sequence 362, App
30	118	3.9	40	10	US-09-864-761-14023	Sequence 44023, A
31	117.5	3.8	522	9	US-08-965-272-2	Sequence 2, Appl1
32	117.5	3.8	659	9	US-10-045-202-4	Sequence 4, Appl1
33	117.5	3.8	945	8	US-08-965-272-2	Sequence 2, Appl1
34	117.5	3.8	2507	9	US-09-819-104A-2	Sequence 2, Appl1
35	117.5	3.8	372	9	US-10-029-180-10	Sequence 10, Appl1
36	117	3.8	543	9	US-09-977-260-14	Sequence 14, Appl1
37	117	3.8	543	9	US-09-977-261-14	Sequence 14, Appl1
38	117	3.8	543	10	US-09-977-269-14	Sequence 14, Appl1
39	117	3.8	543	10	US-09-801-368-140	Sequence 10, App
40	117	3.8	1302	9	US-09-977-260-13	Sequence 13, Appl1
41	116.5	3.8	536	9	US-09-929-266-10	Sequence 10, Appl1
42	116.5	3.8	536	9	US-09-977-261-13	Sequence 13, Appl1
43	116.5	3.8	536	10	US-09-977-269-13	Sequence 13, Appl1
44	116.5	3.8	2005	10	US-09-735-367B-3	Sequence 3, Appl1
45	116.5	3.8				

ALIGNMENTS

RESULT 1
US-09-966-955A-2
Sequence 2, Application US/09966955A
Patent No. US20020155563A1
GENERAL INFORMATION:
APPLICANT: Perez-Villar, Juan J.
APPLICANT: Chang, Han
APPLICANT: Yang, Wen-Pin
APPLICANT: Wu, Yuli
APPLICANT: Whitney, Gena S.
APPLICANT: Kanner, Steven B.
TITLE OF INVENTION: Identification and Cloning of a Full-length Human
TITLE OF INVENTION: Clunk-related Gene, M18F (Mast Cell Immunoreceptor)
TITLE OF INVENTION: Signal Transducer)
FILE REFERENCE: 3053-4113US1
CURRENT APPLICATION NUMBER: US/09/966,955A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/237030
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 443
TYPE: PRT
ORGANISM: HUMAN
FEATURES:
OTHER INFORMATION: HUMAN FULL-LENGTH M18F CDNA CLONE #8, TRANSLATED
OTHER INFORMATION: AMINO ACID SEQUENCE

D	b	1	Meu	sna	arg	in	cl	lys	sn	arg	lys	thr	thr	thr	lys	glu	lys	ser	asn	asp	leu	lys	phe	gln	20
O	y	315	A	AG	CT	CT	CT	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	T	G	374
																								40	
D	b	21	Asn	phe	ser	thr	leu	pro	lys	asn	arg	ser	thr	pro	arg	lys	leu	asn	ser	ala	thr	gln	lys	thr	40
O	y	375	CG	AC	G	GG	T	T	G	G	A	C	CT	C	G	G	A	C	T	C	G	G	A	C	434
																								60	
D	b	41	Gln	arg	met	asn	lys	pro	leu	leu	asn	asp	trp	gln	thr	arg	asn	phe	ala	ala	val	leu	asn	pro	60
O	y	435	GA	AA	A	T	G	C	A	G	T	A	C	A	C	A	C	A	G	T	A	C	A	C	494
																								80	
D	b	61	Ala	lys	gln	lys	thr	ser	asp	asp	trp	trp	arg	asn	pro	gln	leu	arg	met	gln	thr	trp			80
O	y	495	CC	AT	C	A	T	G	AA	A	T	T	T	T	T	T	T	T	T	T	T	T	T	T	554
																								100	
D	b	81	Gln	ser	thr	leu	lys	thr	leu	pro	ala	pro	leu	lys	gln	ser	gln	thr	ala	asp	thr	his			100
O	y	555	T	A	T	T	T	C	A	G	A	T	G	A	T	G	A	G	A	T	G	A	G	A	614
																								120	
D	b	101	Tyr	phe	lys	val	ala	met	asp	thr	pro	leu	pro	leu	asp	thr	arg	thr	ser	ile	ser				120
O	y	615	G	A	G	A	C	A	A	C	C	G	A	T	G	A	G	A	T	G	A	G	A	C	674
																								139	
D	b	121	Gly	asn	pro	thr	trp	asn	thr	gln	--	Thr	arg	leu	gln	arg	val	asp	lys	pro	ile	ser			139
O	y	675	A	A	G	A	T	G	C	A	A	A	C	C	A	A	C	C	T	T	T	T	T	T	734
																								159	
D	b	140	Lys	asn	val	arg	ser	gln	asn	ile	lys	gln	ala	ser	val	arg	lys	asn	lys	ile	pro				159
O	y	735	T	T	G	C	A	C	T	C	T	G	G	C	T	G	C	T	A	T	C	A	C	T	794
																								178	
D	b	160	Leu	pro	pro	pro	ala	arg	pro	leu	ile	thr	leu	pro	lys	lys	thr	gln	pro	leu	pro	--			178
O	y	795	CC	AC	A	G	A	G	A	G	A	C	A	G	A										

Qy	1389	AATGAGTGTTCATTCGTGGAGAGCATTCATGACATACATATATTTCCATCTCG	1448
			:::
Db	379	AspIuLySPhespsrValGuaSpIleIleGluHStyLyasnPheIroIle	398
Qy	1449	CTATAGATGGAGAAAGCAG--GCTGCACCCAGGAACATGCTACCTGACCCAGCCA	1505
Db	399	LeuIleAspIySAspLyStmrgLyValHisAtrGlyGInCysHisLeuThroIinPro	418
Qy	1506	CTGCCTCGCCAGAGCTCTTCTC	1529
Db	419	LeuProIeuThmArgHisLeu	426

RESULT 3
US-09-966-955A-6
; Sequence 6, Application US/09966955A
Patent No. US2002015563A1

APPLICANT: Perez-Villar, Juan J.

APPLICANT: Yang, Wen-ping

APPLICANT: WU, Yuli

APPLICANT: Whitney, Gene

TITLE OF INVENTION: Ident

TITLE OF INVENTION:	Clinical
TITLE OF INVENTION:	
TITLE OF INVENTION:	Sigra

FILE REFERENCE: 3053-4113

CURRENT FILING DATE: 200

; PRIOR APPLICATION NUMBER: 00000000

NUMBER OF SEO ID NOS: 52

```

; SOFTWARE: PatentIn Ver. 2

```

```

; SEQ ID NO 6
;
; LENGTH: 353

```

TYPE: PRT

ORGANISM: HUMAN

OTHER INFORMATION: HUMAN

OTHER INFORMATION: AMINO
; ITS-09-966-955A-6

5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528

Alignment Scores: 1.00
Pred NO .

```

file: no.: 110
Score: 110

```

Percent Similarity:	73.3
Post-Local Similarity:	64.3

```
Best Local Similarity: 04.
Query Match: 36.
```

DB: 9

US-09-856-061-1 (1-1721) x

Our 486 DAGGERCATAAT

[illegible]

Db 3 GlutHrTrpGlnSerI

546 GATACACGCTATTCCA

```

      |||||  |||||:::
db    23 asntbrnistvrpholy

```

DO 23 Aspirin usually

QY 606 GTCTCCACTGAGAGACA

```

Db      43  ILeSerILeGlyGlnPr

```

[illegible]

QY 000 C C I H C C I C H A G G A I O I
 III I I I I I I I I I I

Db 62 ProileSerLysaspva

726 AAGACTCCTTGGCCACC
QY

03 T:CT]aProf ouProPr

DB 02 LYSIIEPIOLEUFI0FI

```

; OTHER INFORMATION: HUMAN M187 SPICE VARIANT CLONE #12, TRANSLATED
; OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-9554-6

Alignment Scores:
Pred. No.: 1,07e-78 Length: 353
Score: 1109.50 Matches: 226
Percent Similarity: 73.22% Conservative: 31
Best Local Similarity: 64.39% Mismatches: 89
Query Match: 36.29% Indels: 5
DB: Gaps: 4

US-09-856-061-1 (1-1721) x US-09-966-9554-6 (1-353)

QY 486 AAGCATGCCCATCAATGAAATTTTACCACCAAGACCTATCCAGATCGAATGGAATACGCA 545
      :::: ||| |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
DB 3 GATHTTPTGTLSTerLysLLeuLeuProLalArgProLLeuLysGlnSerGlnTyrAla 22

QY 546 GATACACGCTATTTCACGAGATATGATGAGACGTCGCCCTTGCTGTACTCCCAAGCGTTCN 605
      ||||| |||||::: ||::: ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 23 ASPTTHHistyPhelysValAlaMetAspHrProLeuProLeuAspHrArgTlnSer 42

QY 606 GTCCTCAGTGAAGCAACACGSGGATGTGAGATGACACAGCTGGGAAGATGGACAA 665
      ::::|::: ||| :::: ||| :::: ||| :::: ||| :::: ||| :::: |||
DB 43 lIeSerlIeGlYlHrProHrTyrPAsnHrGln--ThraGlnLeuGlnAryValAspLys 61

QY 666 CCGTACTTAAAGAGTGTGAGGCCACCGTTTAAAGATTCAAATACACAAAAATTAAC 725
      ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 62 ProLlSerLysSpAlaArgSerLlnslnlIeLysGlnSpAlaSerValArgLysAsn 81

QY 726 AAGACTCTCTTGCCACCTCTCGGCGCTGCTATCACTCTCCCAAGAAGATCAACCCCTTA 785
      ||| ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 82 LyslIleProLeuProProLalArgProLeuLleHrLeuProLysLysTyrGlnProLeu 101

```

RESULT 4
 US-09-864-761-38652
 : Sequence 38652, Application US/09864761
 : Patent No. US20020048763A1
 : GENERAL INFORMATION:
 : APPLICANT: Penn, Sharon G.
 : APPLICANT: Rank, David R.
 : APPLICANT: Hanzel, David K.
 : APPLICANT: Chen, Wensheng
 : TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
 : TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
 : FILE REFERENCE: Aecmlca-X-1
 : CURRENT APPLICATION NUMBER: US/09/864,761
 : CURRENT FILING DATE: 2001-05-23
 : PRIOR APPLICATION NUMBER: US 60/180,312
 : PRIOR FILING DATE: 2000-02-04
 : PRIOR APPLICATION NUMBER: US 60/207,456
 : PRIOR FILING DATE: 2000-05-26
 : PRIOR APPLICATION NUMBER: US 09/632,366
 : PRIOR FILING DATE: 2000-08-03
 : PRIOR APPLICATION NUMBER: GB 24263.6

```

1 LENGTH: 40
2 TYPE: PRF
3 ORGANISM: Homo sapiens
4 FEATURE:
5 OTHER INFORMATION: MAP TO AC005599.5
6 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
7 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
8 OTHER INFORMATION: EXPRESSED IN FETL LIVER, SIGNAL = 5.6
9 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
10 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.4
11 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.7
12 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
13 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4
14 OTHER INFORMATION: SWISSPROT HIT: Q60787, EVALU9 9.00e-10
15
16 US-09-864-761-38652

```

QY	1239	GATGGACATCTTTTGGCCAGCACTCTGTACAAATCCAAAGCGAAGCAACCAATATTTTG	1239
		:::	
Db	1	Asp1ySerPheLeuValAlaArgAspCysSerThrLysSerLysGluGlnIleProTyrValLeu	20
QY	1299	GTGGTCTTTTATGGGAACAGCGCTCATGTGAAATCCGTTTCTCTGAGAGCAATCA	1358
		:::	
Db	21	AlaValaPheIyrLysLysValIyrAsnValLysIleIaTyrPheLeuGluArgAsnGln	40
QY	1359	CAGTTTGCCCTGGGCACA	1376
		:::	
Db	41	GlnPheAlaLeuGlnLyrThr	46

RESULT 5
US-09-764-868-764
; Sequence 764, Application US/09764868


```

US-09-856-061-1 (1-1721) x US-10-186-399-3 (1-675)
OY 750 CCTGTATCATCTCTCCCAAGAGTACCAACCTTATCCSSCCAGCACCAGAGAGAGC 809
DB 141 ProGlyCysThrLeuTrpGluAlaValAsnLeuHisThrAlaValAsnGluGluLys 160
OY 810 AGTGCATCTTCCGTCGCAAGCCCTTCCAGAA---GTCCAGAGGGGGCCAGG--- 863
DB 161 HisArgVal-----ProThrPheProAspArgValLeuLysIleProArgAla 176
OY 864 -----CAGAGAGTGCAGAAAGACTCTAGTAGGGCTCTTGAGCA----- 902
DB 177 ValProValLeuLysMetAspAlaProSerSerSerThrThrLeuAlaGlnTyraAspAsn 196
OY 903 GAAGAGATCTTACCCACAGACCAAGCAAGATCTTCTTCCCATCATCAACCAAAAC 962
DB 197 GluSerLysLysAsnTyrlYserGlnProProSerSerSerThrSerLeuAlaGlnTyrl 216
OY 963 ACACAGAGAGTCCACCTCCATTCGCAAGCTCT-----TCTACATGCGCA 1007
DB 217 AspSerAsnSerLysLysIleTyrlYserGlnProAsnPheAsnMetGlnTyrlIlePro 236
OY 1008 GGA-----AAGCAGAGATATACAAACCCAGA 1031
DB 237 ArgGluAspPheProAspTrpTrpGlnValArgLysLeuLysSerSerSerSerGlu 256
OY 1032 GACCATACAGTACGATGACGACACTGTCTGCTCAGAGATGCCAGAGCTGCAGCCAGC 1091
DB 257 AspValAlaSerSerAsnGlnLys-----GluArgAsnValAsnHis 270
OY 1092 AGCCCTCGAATCTGCTCCCTTGAACAAACAACTCGAGAAACCTGAGCCCAAAAGCTT 1151
DB 271 ThrThrSerLysIleSerTrpGlnPheProGluSerSerSerSerGlu----- 286
OY 1152 GATGAGAGAGTGTGCGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1211
DB 287 GluGluGluAsnLeuAspAspTyrlAspTrpPheAlaGluAsnIleSerArgSerGlnSer 306
OY 1212 GAAGATGTGTTAAATGAAGAGAAAGAGATGTTTGTGTCGAGAGTGTCTTACA 1271
DB 307 GluGlnLeuLeuArgGlnLysGlyLysGlnGluAlaPheMetValArgAsn----- 323
OY 1272 AAATCCAGGCGAAGCAATATGTTTGTGCTGTT-----TATGGAGCAAGCTC 1322
DB 324 SerSerGlnValGlyMetTyrlThrValSerLeuPheSerLysAlaValAsnAspLysLys 343
OY 1323 TACAAATGGAATCCGTTCTCGAGAGACATCAACAGCTTGCCCTGGGCGACAGACTA 1382
DB 344 GlyThrValLysHisTyrlHisValHisThrAsnAlaGluAsnLysLeu-----TyrlLeu 361
OY 1383 CGAGGAATGAGATGTTGATGTTGTGTGAGACATCATTTGAACATCACATAT 1436
DB 362 AlaGluAsnTyrlCysPheAspSerLleProLysLeuIleHisTyrlHisGlnHis 379

RESULT 7
US-09-977-260-4
; Sequence 4, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULTRICH, AXEL
; APPLICANT: GISHITZKY, MIKHAIL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT APPLICATION NUMBER: US/09/977,260
; PRIOR FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545
; PRIOR FILING DATE: 1994-04-22
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 675
; TYPE: PRT

```

```

; ORGANISM: Unknown Organism
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
; OTHER INFORMATION: kinase 2
US-09-977-260-4

Alignment Scores:
Pred. No.: 0.0368 Length: 675
Score: 133.00 Matches: 62
Percent Similarity: 39.92% Conservative: 41
Best Local Similarity: 24.03% Mismatches: 107
Query Match: 4.35% Indels: 48
DB: 9 Gaps: 11

US-09-856-061-1 (1-1721) x US-09-977-260-4 (1-675)
OY 750 CCTGTATCATCTCTCCCAAGAGTACCAACCTTATCCSSCCAGCACCAGAGAGAGC 809
DB 141 ProGlyCysThrLeuTrpGluAlaValAsnLeuHisThrAlaValAsnGluGluLys 160
OY 810 AGTGCATCTTCCGTCGCAAGCCCTTCCAGAA---GTCCAGAGGGGGCCAGG--- 863
DB 161 HisArgVal-----ProThrPheProAspArgValLeuLysIleProArgAla 176
OY 864 -----CAGAGAGTGCAGAAAGACTCTAGTAGGGCTCTTGAGCA----- 902
DB 177 ValProValLeuLysMetAspAlaProSerSerSerThrThrLeuAlaGlnTyraAspAsn 196
OY 903 GAAGAGATCTTACCCACAGACCAAGCAAGATCTTCTTCCCATCATCAACCAAAAC 962
DB 197 GluSerLysLysAsnTyrlYserGlnProProSerSerSerThrSerLeuAlaGlnTyrl 216
OY 963 ACACAGAGAGTCCACCTCCATTCGCAAGCTCT-----TCTACATGCGCA 1007
DB 217 AspSerAsnSerLysLysIleTyrlYserGlnProAsnPheAsnMetGlnTyrlIlePro 236
OY 1008 GGA-----AAGCAGAGATATACAAACCCAGA 1031
DB 237 ArgGluAspPheProAspTrpTrpGlnValArgLysLeuLysSerSerSerSerGlu 256
OY 1032 GACCATACAGTACGATGACGACACTGTCTGCTCAGAGATGCCAGAGCTGCAGCCAGC 1091
DB 257 AspValAlaSerSerAsnGlnLys-----GluArgAsnValAsnHis 270
OY 1092 AGCCCTCGAATCTGCTCCCTTGAACAAACAACTCGAGAAACCTGAGCCCAAAAGCTT 1151
DB 271 ThrThrSerLysIleSerTrpGlnPheProGluSerSerSerSerGlu----- 286
OY 1152 GATGAGAGAGTGTGCGAGAAATGATGATGATGATGATGATGATGATGATGATGATG 1211
DB 287 GluGluGluAsnLeuAspAspTyrlAspTrpPheAlaGluAsnIleSerArgSerGlnSer 306
OY 1212 GAAGATGTGTTAAATGAAGAGAAAGAGATGTTTGTGTCGAGAGTGTCTTACA 1271
DB 307 GluGlnLeuLeuArgGlnLysGlyLysGlnGluAlaPheMetValArgAsn----- 323
OY 1272 AAATCCAGGCGAAGCAATATGTTTGTGCTGTT-----TATGGAGCAAGCTC 1322
DB 324 SerSerGlnValGlyMetTyrlThrValSerLeuPheSerLysAlaValAsnAspLysLys 343
OY 1323 TACAAATGGAATCCGTTCTCGAGAGACATCAACAGCTTGCCCTGGGCGACAGACTA 1382
DB 344 GlyThrValLysHisTyrlHisValHisThrAsnAlaGluAsnLysLeu-----TyrlLeu 361
OY 1383 CGAGGAATGAGATGTTGATGTTGTGTGAGACATCATTTGAACATCACATAT 1436
DB 362 AlaGluAsnTyrlCysPheAspSerLleProLysLeuIleHisTyrlHisGlnHis 379

RESULT 8
US-09-977-261-4
; Sequence 4, Application US/09977261
; Publication No. US20030054527A1
; GENERAL INFORMATION:

```

```

APPLICANT: ULLRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
APPLICANT: SURES, IRMINGARD
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1259
CURRENT APPLICATION NUMBER: US/09/977,261
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232,545
PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 675
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
US-09-977-261-4

```

```

Alignment Scores:
Pred. No.: 0.0368 Length: 675
Score: 133.00 Matches: 62
Percent Similarity: 39.92% Conservative: 41
Best Local Similarity: 24.03% Mismatches: 107
Query Match: 4.35% Indels: 48
DB: Gaps: 11
US-09-856-061-1 (1-1721) x US-09-977-261-4 (1-675)

```

```

QY 750 CCGTGATCTACTCTCCCAAGAGTACCAACCTTACCCAGACACAGAGAGAGC 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 141 ProglycylserthrleuprglnalatyralasnleuHistrAlavalasnlguglulys 160
QY 810 AGTCATCTCTCGCTCCCAAGCCCTTCCAGAA---GTCCAGAGGGGGCCAGG--- 863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 161 HisArgVal-----ProthrPheproaspArgValleuylsIleProArgAla 176
QY 864 -----CAGAGAGTGCAGAAAGACTTCAGTAGGGTCCCTGGAGCA----- 902
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 ValProValleuylsMetaspAlaProserSerSerthrThrleuAlagIntyrAspAsn 196
QY 903 GAAGAGAATCTCACACAGCAAGCAAGCAAGATCTTCTGCCCATCATCAACCAAAAC 962
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 GluberlyslsAsnTyrIlglySerGlnProProserSerSerThrSerleuAlagIntyr 216
QY 963 ACACAGAGAGTCCACCTGCCATTGCCAGCTCT-----TCTTACATGCCA 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 217 AspSerAsnSerlyslsIleTyrGlySerGlnProAsnPhasnMetGlnTyrIlePro 236
QY 1008 GGA-----AAGCAGATATACAGCCAGA 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 237 ArgGluAspPheProAspTyrIlePrlnValArgylsleuylsSerSerSerSerGlu 256
QY 1032 GACCATACAGGTAGCATGACACACTGCTCCTCAGAGATGCCAGTGCAGCCAGCCAC 1091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 257 AspValAlaSerSerAsnGlnLys-----GluArgAsnValAlasnHis 270
QY 1092 AGCCCTGCAATGCTGCCCTATGAAAACAACAACCTCGGAGAAACCTGCAGCCACAAGCCT 1151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 271 ThrThrSerlyslsIleSerTyrIlePrlnheProGluSerSerSerSerGlu----- 286
QY 1152 GATGAGAGAGATGTCTGCAGAGATAGTATGATGAGATATACAGTGCAGCCAGCATG 1211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 287 GluGluGluAsnLeuAspAspTyrAspTyrPheAlaGlnIleSerArgSerGlnSer 306
QY 1212 GAAGATGCTTATATGAAGACAAGAGATGCTATTTTGGTCCGAGACTGCTCTACA 1271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 307 GluGlnLeuLeuArgGlnIlysglySerGlnIlyAlaPheMetValArgAsn----- 323
QY 1272 AATTCAGAGGAGCAACATATGTTTGGAGTGT-----TATGGCAACAGGTC 1322
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 324 SerSerGlnValIglyMetIlyThrValSerLeuPheSerlyslsAlaValAlasnAspIlysls 343

```

```

QY 1323 TACATGTGAATCCGTTTCTCGAGCAATCAACAGTTTGCCCTGGCAGAGACTA 1382
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 344 GlyThrValIlyshIstYrHisValHisThrAsnAlaGluAsnLysLeu-----TyrIleu 361
QY 1383 CGAGGAATGAGATGTGATTCGTGCGAAGACATCATTTGAACACTACATAT 1436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 362 AlaGluAsnTyrGlyCysPheAspSerIleProIlyslsleuIleHstYrHisGlnHis 379

```

```

RESULT 9
US-09-977-269-4
Sequence 4, Application US/09977269
Patent No. US20020082037A1
GENERAL INFORMATION:
APPLICANT: ULLRICH, AXEL
APPLICANT: GISHIZKY, MIKHAIL
APPLICANT: SURES, IRMINGARD
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
FILE REFERENCE: 038602/1260
CURRENT APPLICATION NUMBER: US/09/977,269
CURRENT FILING DATE: 2001-10-16
PRIOR APPLICATION NUMBER: 08/232,545
PRIOR FILING DATE: 1994-04-22
NUMBER OF SEQ ID NOS: 24
SOFTWARE: Patentln Ver. 2.1
SEQ ID NO 4
LENGTH: 675
TYPE: PRT
ORGANISM: Unknown Organism
FEATURE:
OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
US-09-977-269-4

```

```

Alignment Scores:
Pred. No.: 0.0368 Length: 675
Score: 133.00 Matches: 62
Percent Similarity: 39.92% Conservative: 41
Best Local Similarity: 24.03% Mismatches: 107
Query Match: 4.35% Indels: 48
DB: Gaps: 11
US-09-856-061-1 (1-1721) x US-09-977-269-4 (1-675)
QY 750 CCGTGATCTACTCTCCCAAGAGTACCAACCTTACCCAGACACAGAGAGAGC 809
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 141 ProglycylserthrleuprglnalatyralasnleuHistrAlavalasnlguglulys 160
QY 810 AGTCATCTCTCGCTCCCAAGCCCTTCCAGAA---GTCCAGAGGGGGCCAGG--- 863
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 161 HisArgVal-----ProthrPheproaspArgValleuylsIleProArgAla 176
QY 864 -----CAGAGAGTGCAGAAAGACTTCAGTAGGGTCCCTGGAGCA----- 902
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 ValProValleuylsMetaspAlaProserSerSerthrThrleuAlagIntyrAspAsn 196
QY 903 GAAGAGAATCTCACACAGCAAGCAAGCAAGATCTTCTGCCCATCATCAACCAAAAC 962
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 GluberlyslsAsnTyrIlglySerGlnProProserSerSerThrSerleuAlagIntyr 216
QY 963 ACACAGAGAGTCCACCTGCCATTGCCAGCTCT-----TCTTACATGCCA 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 217 AspSerAsnSerlyslsIleTyrGlySerGlnProAsnPhasnMetGlnTyrIlePro 236
QY 1008 GGA-----AAGCAGATATACAGCCAGA 1031
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 237 ArgGluAspPheProAspTyrIlePrlnValArgylsleuylsSerSerSerSerGlu 256
QY 1032 GACCATACAGGTAGCATGACACACTGCTCCTCAGAGATGCCAGTGCAGCCAGCCAC 1091
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 257 AspValAlaSerSerAsnGlnLys-----GluArgAsnValAlasnHis 270
QY 1092 AGCCCTGCAATGCTGCCCTATGAAAACAACAACCTCGGAGAAACCTGCAGCCACAAGCCT 1151
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 271 ThrThrSerLysLeuSerTrpGluPheProGluSerSerSerSerGlu----- 286
QY 1152 GATGAGAGAGATGCTGGCAGCAATGTAATGTACATTGGAGATACAGTCCCGCAGCAGTg 1211
Db 287 GUGUGLUGLUNASmLeuAspAspTrpPheAlaGlyAsnLeuSerSerGlnSer 306
QY 1212 GAAGATGCTGTAATGAAGAGAACAGAGATGTAATTTTGGTCCGAGACTGCTTACA 1271
Db 307 GUGUGLUGLUNASmLeuAspAspTrpPheAlaGlyAsnLeuSerSerGlnSer 323
QY 1272 AATTCAGAGCAGAACCATATGTTGTTGGTGTGTT-----TATGGAGAACAGTgC 1322
Db 324 SerSerGlnValGlyMetLeuTrpValSerLeuPheSerLysAlaValAsnAspLysLys 343
QY 1323 TACAAATGTAATCCGTTTCCCGAGAGCAATCAACAGATTGCGCTGGCAGCAGACTA 1382
Db 344 GLyThValLysHisLysTrpHisValHisLysHisLysHisLysHisLysHisLys 361
QY 1383 CCAGGAATGAGATGTTGATTCTGTGGAAGACATCATGGAACACTACATAT 1436
Db 362 AlAGLUNASmTrpCysPheAspSerLeuProLysLeuHisLysHisLysHis 379

RESULT 10
US-10-081-980B-1
; Sequence 1, Application US/10081980B
; Publication No. US2003004137A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Gibbs, E. Michael
; APPLICANT: McNeilsh, John D.
; TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form Of
; FILE REFERENCE: Subunit Of PI-3 Kinase
; CURRENT APPLICATION NUMBER: US/10/081.980B
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/270,014
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 688
; TYPE: PRT
; ORGANISM: mus musculus
US-10-081-980B-1

Alignment Scores:
Pred. No.: 0.037 Length: 688
Score: 133.00 Matches: 56
Percent Similarity: 46.86% Conservative: 26
Best Local Similarity: 32.00% Mismatches: 75
Query Match: 4.35% Indels: 19
Gaps: 9

US-09-856-061-1 (1-1721) x US-10-081-980B-1 (1-688)
QY 913 CTCACCAACGACAGAACGCAATCTTCTT---GCCATCATCAACCAACACACAGAGA 969
Db 250 LeuSerGlnAlaSerSerLysAsnLeuAsnAlaArgValLeuSerGluLeuPheSer 269
QY 970 AGAGTCACCTGCGATTGTCAGCTCTTCTTCAATGCGCAGAGCAAGATATCAAGCA 1029
Db 270 ProValLeuPheArgPheProAla-----AlaSerSerAspAsnTrpGluHisLeuLe 287
QY 1030 GAGACCATACAGATGATGATGACGACACTGCTCTCAGAGATGCCAGCTGACGCCAGCC 1089
Db 288 LysAlaIleGluLeuLeuLeuSerThr-----GluTrpAsnGlnArgGlnProAla 304
QY 1090 ACAGCCCTCGAATGCTGCCCTATGAACACAACTCGGAGAACCTGAGCCCAAGC 1149
Db 305 ProAlaLeuProProLysProProLysProThrThrValAlaAsn-----AsnSer 321
QY 1150 CTGATGAGAAGATGCTGTG---CAGAAATGAATGTATGAGAAATACAGTCCGAG 1205

```

```

Db 322 MetAsnAsnAsnMetSerLeuGlnAspAlaGluTrpTrpGlyAspLeuSerArgL 341
QY 1206 GCAGTGGAGATGCTGTTAATGAAGAGAACAGATGGTACTTTTGGTCCGAGACTGC 1265
Db 342 GluValAsnGlnLysLeu-----ArgAspThrAlaAspGlyThrPheLeuValArgAspAla 360
QY 1266 TCTACAAAATCCAGAGCAGCAACCATATGTTTGGTGTGTTTATGAGAACAGCTAC 1325
Db 361 SerThrLysMetHisGlyAsp---TyrThrLeuThrLeuArgLysGlyGlyAsn----- 377
QY 1326 AATGTGAAGAAATCCGTTTCCCTGAGAGCAATCAACAGTTTGCCCTGGCAGCAGACTACGA 1385
Db 378 AsnLysLeuLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 396
QY 1386 GGAATGAGATGTTGATTCTGTGGAACACATCATGGAACCTGAC 1430
Db 397 -----ThrPheAsnSerValValGluLeuLeuAsnHisLys 408

RESULT 11
US-10-081-980B-3
; Sequence 3, Application US/10081980B
; Publication No. US2003004137A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Gibbs, E. Michael
; APPLICANT: McNeilsh, John D.
; TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form
; FILE REFERENCE: Subunit Of PI-3 Kinase
; CURRENT APPLICATION NUMBER: US/10/081.980B
; PRIOR FILING DATE: 2002-08-23
; PRIOR APPLICATION NUMBER: 60/270,014
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 724
; TYPE: PRT
; ORGANISM: mus musculus
US-10-081-980B-3

Alignment Scores:
Pred. No.: 0.0375 Length: 724
Score: 133.00 Matches: 56
Percent Similarity: 46.86% Conservative: 26
Best Local Similarity: 32.00% Mismatches: 75
Query Match: 4.35% Indels: 19
Gaps: 9

US-09-856-061-1 (1-1721) x US-10-081-980B-3 (1-724)
QY 913 CTCACCAACGACAGAACGCAATCTTCTT---GCCATCATCAACCAACACACAGAGA 969
Db 250 LeuSerGlnAlaSerSerLysAsnLeuAsnAlaArgValLeuSerGluLeuPheSer 269
QY 970 AGAGTCACCTGCGATTGTCAGCTCTTCTTCAATGCGCAGAGCAAGATATCAAGCA 1029
Db 270 ProValLeuPheArgPheProAla-----AlaSerSerAspAsnTrpGluHisLeuLe 287
QY 1030 GAGACCATACAGATGATGATGACGACACTGCTCTCAGAGATGCCAGCTGACGCCAGCC 1089
Db 288 LysAlaIleGluLeuLeuLeuSerThr-----GluTrpAsnGlnArgGlnProAla 304
QY 1090 ACAGCCCTCGAATGCTGCCCTATGAACACAACTCGGAGAACCTGAGCCCAAGC 1149
Db 305 ProAlaLeuProProLysProProLysProThrThrValAlaAsn-----AsnSer 321
QY 1150 CTGATGAGAAGATGCTGTG---CAGAAATGAATGTATGAGAAATACAGTCCGAG 1205
Db 322 MetAsnAsnAsnMetSerLeuGlnAspAlaGluTrpTrpGlyAspLeuSerArgL 341
QY 1206 GCAGTGGAGATGCTGTTAATGAAGAGAACAGATGGTACTTTTGGTCCGAGACTGC 1265

```

```

Db      342 GLUValAsnGluLysLeu---ArgAspThrAlaAspGlyThrPheLeuValAlaGAspAla 360
OY      1266 TCACAAATCCAGAGCAGACATATGTTGGTGGTGGTTTATGGAAACAAGTCTAC 1325
Db      361 SerThrLysMetHisGlyAsp---TyrThrLeuThrLeuArgLysGlyGln----- 377
OY      1326 AATGTGAAATCCGTTTCCTCGAGACATCAACAGTTGGCCCTGGGACAGACTACGA 1385
Db      378 AsnLysLeuLeuLysLysPheHisArgAspGlyLysTyrGlyPheSerAspProLeu--- 396
OY      1386 GGAATGAGATGTTGTGTTCTGTGGAAGACATCATGACATGACATAC 1430
Db      397 -----ThrPheAsnSerValValGluLeuLeuLeuHisIstYr 408

RESULT 12
US-09-962-929-4
; Sequence 4, Application US/09962929
; Patent No. US20020115058A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.
; TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated with A
; TITLE OF INVENTION: Expression of Plk3r1
; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,929
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/668,644
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-962-929-4

Alignment Scores:
Pred. No.: 0.0376 Length: 724
Score: 133.00 Matches: 96
Percent Similarity: 37.32% Conservative: 57
Best Local Similarity: 23.41% Mismatches: 169
Query Match: 4.35% Indels: 89
Gaps: 16

US-09-856-061-1 (1-1721) x US-09-962-929-4 (1-724)
OY      371 GTGTGACGCGGTTCT-----GGAACCACTTCGGATGCACAG-----AAG 409
Db      32 ValAsnLysGlySerLeuValAlaLeuLysPheSerAspGlyGlnGluAlaArgProGlu 51
OY      410 GAAGTGGCTGGGGTCCAGGT-----GGAGAAAATGCAACAGTAAACA 454
Db      52 GlnIleGlyTyr-LeuAsnGlyTyrAsnGluThrThrGlyGluArgGlyAspPheProGlu 71
OY      455 GCAGTACGAAGATCCTGAGTCCAGCTGCTGTAAGGATGCGCATGAAGAATTTTAC 514
Db      71 YThrYrValGluYrIleGlyArgLysLysLysSerProPheProIleProLysProArgPr 91
OY      515 AGCCAGACCTATCCAGATCGAATCGAGATACAGCATATTCGAGATATGATGA 574
Db      91 OProArgProLeuProValAlaPro---GlySerSerLysThrGluAlaAspValGluGlu 110
OY      575 GGCTCCCTCTGTACTCTCCCAAGGCTTCT----- 605
Db      110 nGlnAlaLeuThrLeuProAspLeuAlaGluGlnPheAlaProAspIleAlaProPr 130
OY      606 -----GT 607
Db      130 oLeuLeuIleLysLeuValGluAlaIleGluLysLysGlyLeuGluLysSerThrLeuTy 150
OY      608 CTCACACTGAGAGCAAAACCGAGGATGTGAGATGACACAGCTGGAAGAAGTGACACGCC 667

```

```

Db      150 rArgThrGlnSerSerSerAsnLeuAlaGluLeuArgGlnLeuLeuAspCysAspThrPr 170
OY      668 TACCTTCAG-----GATCTCAAGAACCAACCTTTAAAGATTCAA 709
Db      170 oSerValAspLeuGluMetIleAspValHisValIleuAlaAspAlaPheLysArgTyrLe 190
OY      710 ATACACAAAATTAACAAACATCCTTTGGCCACCTCTCGGCTGCTATCACTCTCCCA 769
Db      190 u-----LeuAspLeuProAsnProValIleProAlaAlaVala 202
OY      770 GAAGTCAACACCTTACCCACACACAGACAGAGAGAGAGAGTGCATCTGCTCCA-- 827
Db      202 IYrSerGluMetIleSerLeuAlaProGluValGlnSerSerGluTyrIleGlnLe 222
OY      828 -----AAGCCACCTTTCAGAAATTCACAGAGGGGCCACAGCAGNG 868
Db      222 uLeuLysLysLeuIleArgSerProSerIleProHisGlnTyrTrpLeuThrLeuGlnTyr 242
OY      869 GAGTGCAAAAGACTTCAGTAGGGCTCCTTGAGCAGAAAGAAATCTCACACACAGACAAA 928
Db      242 rLeuLeuLysHisPhePheLysLeu-----SerGlnThrSe 254
OY      929 GCCAGATCTTCTGGCCATCATCAACCAAAACACACAGAGATGCACCTGCCATTGC 988
Db      254 rSerLysAsnLeuLeuAsnAlaArgValLeuSerGluIlePheSerProMetLeuPheAr 274
OY      989 CAGCTCTCTCATATGCCAGAGAAACAGCATATACAGCAGACCATATACAGTATGACAT 1048
Db      274 gPheSerAla-AlaSerSerAspAsnThrGluAsnLeuIleLysValIleGluIleLeuI 294
OY      1049 GCAGCACTGTCTGCTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1108
Db      294 lSerThr-----GluTrpAsnGluArgGlnProAlaProAlaLeuProProLysP 311
OY      1109 CTATGAAACCAAACTCGGAGAAACCTGACCCACAAACCTGATGAGAGATGCTG 1168
Db      311 rProLysProThrThrValAlaAsn-----AsnGlyMetAsnAsnMetSerI 328
OY      1169 G-CAGAT-----GAATGTATATGAGATATACAGTCCGAGGAGGATGATGTTAA 1224
Db      328 eGlnAsnAlaGluTrpTyrTrpLysPheLysPheSerHisGlyGlnGluValAsnGluLysLeu 347
OY      1225 TGAAGAGAACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1284
Db      348 --ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGly 367
OY      1285 AACCATATGTTTGTGTGTTTATGGAACAGTCTACAAATGTAATGTAATGCTTCC 1344
Db      367 sp---TyrThrLeuThrLeuArgLysGlyGln-----AsnLysLeuIleLysLysP 384
OY      1345 TCGAGAGCAATCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1404
Db      384 hHisArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheSer 400
OY      1405 CTGTGAGACATCATTCACATGACATGACATGACATGACATGACATGACATGACATGAC 1430
Db      400 eValValGluLeuIleLeuHisIstYr 408

RESULT 13
US-09-962-929-2
; Sequence 2, Application US/09962929
; Patent No. US20020115058A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.
; TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated wit
; TITLE OF INVENTION: Expression of Plk3r1
; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,929
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/668,644

```

```

1  PRIOR APPLICATION NUMBER: US 09/632,366
2  PRIOR FILING DATE: 2000-08-03
3  PRIOR APPLICATION NUMBER: GB 24263.6
4  PRIOR FILING DATE: 2000-10-04
5  PRIOR APPLICATION NUMBER: US 60/236,359
6  PRIOR FILING DATE: 2000-09-27
7  PRIOR APPLICATION NUMBER: PCT/US01/00666
8  PRIOR FILING DATE: 2001-01-30
9  PRIOR APPLICATION NUMBER: PCT/US01/00667
10 PRIOR FILING DATE: 2001-01-30
11 PRIOR APPLICATION NUMBER: PCT/US01/00664
12 PRIOR FILING DATE: 2001-01-30
13 PRIOR APPLICATION NUMBER: PCT/US01/00669
14 PRIOR FILING DATE: 2001-01-30
15 PRIOR APPLICATION NUMBER: PCT/US01/00665
16 PRIOR FILING DATE: 2001-01-30
17 PRIOR APPLICATION NUMBER: PCT/US01/00668
18 PRIOR FILING DATE: 2001-01-30
19 PRIOR APPLICATION NUMBER: PCT/US01/00663
20 PRIOR FILING DATE: 2001-01-30
21 PRIOR APPLICATION NUMBER: PCT/US01/00662
22 PRIOR FILING DATE: 2001-01-30
23 PRIOR APPLICATION NUMBER: PCT/US01/00661
24 PRIOR FILING DATE: 2001-01-30
25 PRIOR APPLICATION NUMBER: PCT/US01/00670
26 PRIOR FILING DATE: 2001-01-30
27 PRIOR APPLICATION NUMBER: US 60/234,687
28 PRIOR FILING DATE: 2000-09-21
29 PRIOR APPLICATION NUMBER: US 09/608,408
30 PRIOR FILING DATE: 2000-06-30
31 PRIOR APPLICATION NUMBER: US 09/774,203
32 PRIOR FILING DATE: 2001-01-29
33 NUMBER OF SEQ ID NOS: 49117
34 SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
35 SEQ ID NO 34590
36 LENGTH: 511
37 TYPE: PRT
38 ORGANISM: Homo sapiens
39
40 FEATURE:
41 OTHER INFORMATION: MAP TO AB203049.1
42 OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.7
43 OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 9.3
44 OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.9
45 OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
46 OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
47 OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.2
48 OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.5
49 OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.8
50 OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.8
51 OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.1
52 OTHER INFORMATION: SWISSPROT HIT: Q02817, EVALUE 3.00e-24
53 OTHER INFORMATION: EST_HUMAN HIT: AA487852.1, EVALUE 2.00e-73
54 US-09-866-761-34590
55
56 Alignment Scores:
57 Pred. No.: 0.0759 Length: 511
58 Score: 128.50 Matches: 104
59 Percent Similarity: 33.04% Conservative: 45
60 Best Local Similarity: 23.06% Mismatch: 163
61 Query Match: 4.20% Indels: 139
62 DB: 10 Caps: 23
63
64 US-09-856-061-1 (1-1721) x US-09-864-761-34590 (1-511)
65
66 Oy 456 GACTCGAAGATCCTGCAATTCACGCTGCTGAGCATGCGCATCATGAAATTTTACCA 515
67 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
68 2 AspGngInUsErProAsPaLaCySleuPProThrVaLProGluAlaProAlaProPro 21
69
70 Oy 516 GCCAGATTCATCCAGGATGGAATATAGCAGATACAGCGATATTTCCAGGATATGATGAG 575
71 ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
72 22 GlnlyPProLeuAaSnSerGlnSerGln-----LysHsLleuAlaPro 35
73
74 576 GCTCCCTCTGTACTCTCCCAAGGCTTCTGTCTCC---ACTGAGAGACAACACGAGAT 632
75

```



```

Db      36  ProProleuLeuSerProleuLeuProSerilelyProthrValaIgllyThArgIn 55
QY      633  GTGAGATGACACAGCTGAGAAAGTg-----GAC 662
Db      56  AspGlySerIngluAlaProGluAlaProLeuSerSerGluLeuGluProPheHisPro 75
QY      663  AAGCCTTACCTTCAGAGATGTGACAGCCAGCTTTAAAGATTCAATACACAAATA 722
Db      76  LysProLysIleArgThrArgLysSerSerArgMetThrProPheProAlaThr----- 93
QY      723  AACAGATCTCTTGGCCAGCTGCGCCGCTATCACTCCCAAGAGTACCAAGTACCAACC 782
Db      94  -----SerAlaAlaProGluProHisProSerThrSerThrAla-----GlnPro 108
QY      783  TTACCCCGACGACACACAGAGAGAGC-----AGTGCATCTTC 821
Db      109  ValThrProLysProThrSerGlnAlaThrArgSerArgThrAsnArgSerSerValys 128
QY      822  GCTCAAAAGCCACCTTTCa----- 842
Db      129  ThrProGluProValaIalProThrAlaProGluLeuGlnProSerThrSerThrAspIn 148
QY      843  -----GAGTCCAGAGGGGCCCGCCAGGACAGAGAGTGCa--- 875
Db      149  ProValThrSerGluProThrSerGlnAlaThrArgGlyArgLysSerThrSerVal 168
QY      876  AAAGACTTCAGTAGGCTCTTGAGACAGAGAAGAATCTCACCCACAGACa----- 926
Db      169  LysThrProGluThrValaIalProThrAlaLeuGlnProSerThrSerThrAsp 188
QY      927  AAGCCAGATCTCTTGGCCATCATCA-----AACCAAAACCA 965
Db      189  ArgProValThrSerGluProThrSerGlnAlaThrArgGlyArgLysSerThrSer 208
QY      966  CAGAGAGTCA-----CTGCGCATTCGACGCTCTCTACATCCAGAAAGACa----- 1017
Db      209  ValLysThrProGluProValaIalProThrAlaProGluLeuGlnProSerThrSer 228
QY      1018  -----GTATACAAAGCCAGAGACATACAGTACGATCAGACTGCTCTGCTGAG 1067
Db      228  rAspGlnProValThrSerGluProThrArgLysAlaThrArgGlyArgLysAsnArgSe 248
QY      1068  AATGTCAGAGTGC-----CAGCCAGCC-----ACAGCCCTCGAATGCTGCATGAAAC 1118
Db      248  rSerValLysThrProGluProValaIalProThrAlaProGluLeuArgPro-----Ser 267
QY      1119  ACAAACTCGAGAAACCTGACCCACAAAGCTGATGAGAGAGATGCTGCGAGATGAA 1178
Db      267  rSerThrAspArgProValThrProLysProThrSerArgThrThrArgSerArgThrAs 287
QY      1179  TGGTACATTGGAGATACa----- 1197
Db      287  nMetSerSerValLysThrProGluThrValaIalProThrAlaProGluLeuGlnLese 307
QY      1198  -----GTCGCGAGCGACTGGAAGTGTGTTA 1223
Db      307  rThrSerThrAspGlnProValaIalProLysProThrSerArgThrThrArg----- 324
QY      1224  ATGAAAGTGAACAGATGCTACTTTTGTCCGAGAGTCTCTCAAAATCCAAAGCA 1283
Db      325  -----SerArgThrAsnMet-----SerSerValLysAsnProLuse 337
QY      1284  GAACCATATGTTTGGTGTGTTTATGCGAACACAGGCTACATATGAAATCCGTTTC 1343
Db      337  rThrValProIleAlaProGluLeu-----ProProSerThrSerThrGluGlnProVal 356
QY      1344  CTCGAGAGCATCAACAGATTGCTCGGACAGACTACAGAGATGAGATGTTTGAAT 1403
Db      356  hrProGlu-----ProThrSerArgAlaThrArg----- 365
QY      1404  TCTGTGAGACATCATGTAACACTACATATTTTCCATCTGCTAATAGATGGGAAA 1463

```

```

Db      366  -----GlyArgLysAsn-----ArgS 371
QY      1464  GACAGGCTGCAGGACGAGAAACAGTGTCTACCCAGCAGCTGCTCGCAGAGCTC 1523
Db      371  erSerGlyLysThrProGluThrLeuValPro-----ThrAlaProLysLeuGluP 388
QY      1524  CTCTCAGTCACTCAGTACCTCAGCAGGCACT 1552
Db      388  roSerThrSerThrAspGlnProValThr 397

RESULT 15
US-10-186-399-2
: Sequence 2, Application US/10186399
: Patent No. US20020173481A1
: GENERAL INFORMATION:
: APPLICANT: Ekman, Niklas
: APPLICANT: Arighi, Elena
: APPLICANT: Vastrik, Imre
: APPLICANT: Tamagnone, Luca
: APPLICANT: Allitalo, Karl
: TITLE OF INVENTION: REGULATION OF VASCULAR ENDOTHELIUM USING BMX TYROSINE
: FILE REFERENCE: 28113/31941A
: CURRENT APPLICATION NUMBER: US/10/186,399
: PRIOR FILING DATE: 2002-07-01
: PRIORITY APPLICATION NUMBER: US 08/320,432
: NUMBER OF SEQ ID NOS: 5
: SOFTWARE: PatentIn Ver. 2.0
: SEQ ID NO 2
: LENGTH: 651
: TYPE: PRT
: ORGANISM: Mus musculus
US-10-186-399-2

Alignment Scores:
Pred. No.: 0.0896 Length: 651
Score: 128.00 Matches: 61
Percent Similarity: 39.92% Conservative: 40
Best Local Similarity: 24.11% Mismatches: 90
Query Match: 4.19% Indels: 62
Gaps: 11

US-09-856-061-1 (1-1721) x US-10-186-399-2 (1-651)
QY      750  CCTGCTATCACTCCCTCCCAAGATACCAACCTTACCCAGCACACAGAGAGAGC 809
Db      141  ProGlyCysThrLeuTrpGluAlaTrAlaAspLeuHisIleAlaIleSerAspLeuLys 160
QY      810  ACTGCATCTTCGCTCAAAAGCCACCTTTCAGAA-----GTCCAGAGGGGGCCAGC 863
Db      161  HisArg-----AlaProThrPheProGluArgLeuLeuLysIleProArgAla 176
QY      864  -----CAGAGAGTGCAGAAAGACTTCAGTAGGCTCTT----- 896
Db      177  ValProValLeuLysMetAspAlaSerSerSerGlyAlaIleLeuProGlnTyAspSer 196
QY      897  -----GGAGCAGAGAGAAATCTCACACAGACAGCAAGACAGCAAGAAATCT 938
Db      197  TyrSerLysLysSerCysGlySerGlnProThrSerAsnIleArgTyrIleProArgGlu 216
QY      939  TCTTGCCCATCATCAACCA-----AACACACAGAAAGAGTCCAGCTGCAATGCGAGCT 995
Db      217  AspCysProAspTrpTrpIleValaIalGlyLysLeuLysSerGluAlaAspIleAla----- 234
QY      996  TCTTACATGCCAGAAAGCACAGTATACAGCCAGACACATACAGATGACAGACAC 1055
Db      234  ----- 234
QY      1056  TGTCTCTCAGAGATGCCAGCTGACGACGACGACGCTGATGCTGCTAT--- 1112
Db      235  CysSerAsnGlnLeuGlnArgAsnIleAlaSerHisSerThrSerLysMetSerTrpGly 254

```

```
OY 1113 -----GAAACACAACCTCGAGAAACCTGACCCCAAGCCCTGATGAGAGGATGTC 1166
      |||:|||||:|||||
Db 255 PheProGluSerSerSerSerGlu-----GluGluGluAsnLeu 267
      |||:|||||:|||||
OY 1167 TGCAGAAATGAAATGCTACATTGGAGATACAGTCGCCAGGCACTGGAAGATGTGTTAATG 1226
      |||:|||||:|||||
Db 268 HisAlaTyrAspTyrPheAlaGlyAsnIleSerArgSerGlnSerGluGlnLeuLeuArg 287
      |||:|||||:|||||
OY 1227 AAAGAGACAAGATGCTACTTTTGGTCCGAGACTGCTCTACAAATCCAGGCAGAA 1286
      |||:|||||:|||||
Db 288 GlnLysGlyLysGlnGlyAlaPheMetValArgAsn-----SerSerGlnMetGly 304
      |||:|||||:|||||
OY 1287 CCATATGTTTGGTGGTGT-----TATGGACAAGAGTCTACAAATGTGAAAATC 1337
      |||:|||||:|||||
Db 305 MetTyrThrValSerLeuPheSerLysAlaValAsnAspLysGlyThrValLysHis 324
      |||:|||||:|||||
OY 1338 CGTTTCTCGAGAGCAATCACAGCTTTCCTGGGCACAGGACTTACGAGGAATGAGATG 1397
      |||:|||||:|||||
Db 325 TyrHisValHisThrAsnAlaGluAsnLysLeu-----TyrLeuAlaGluAsnTyrCys 342
      |||:|||||:|||||
OY 1398 TTGATTTCTGTGAAGACATCATGAACTACACATAT 1436
      |||:|||||:|||||
Db 343 PheAspSerIleProLysLeuIleHisTyrHisGlnHis 355
      |||:|||||:|||||
```

Search completed: April 21, 2003, 12:35:25
Job time : 58.7798 secs

GenCore version 5.1.4-P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:15:56 ; Search time 53.7435 Seconds
(without alignments)
6156.924 Million cell updates/sec

Title: US-09-856-061-1

Perfect score: 3057
Sequence: 1 acgagcgcccaactgcgccag.....taaaaaaaaaaaaaaaaa 1721

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2.1/USPTO.spool/US09856061/runat_21042003.113019.3758/app.query.fasta_1.3150
-DB=PIR_73 -OEFM=fastan -SUFFIX=n2p.rpr -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFM=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061@cgn2.1.151.0@runat_21042003.113019.3758 -NCPU=6 -ICPU=3
-NO_XLPEX -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TTIMEOUT=120
-MARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

PIR_73:*
1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412.5	13.5	533	2 B56110	tyrosine phosphopr
2	398	13.0	533	2 A56110	tyrosine phosphopr
3	158.5	5.2	663	1 TVMVR	protein-tyrosine k
4	149.5	4.9	572	1 T12496	hypothetical prote
5	148.5	4.9	1029	2 T30351	mucin-like protein
6	146	4.8	728	2 H59435	phosphoinositide-3
7	145	4.7	723	2 B38749	3-phosphatidylinos
8	140	4.6	1733	2 S27939	tensin - chicken
9	140	4.6	1744	2 A54970	tensin, cardiac mu
10	139.5	4.6	1792	2 A57075	tensin - chicken (
11	138.5	4.5	1603	2 S17983	gene posterior sex
12	138.5	4.5	443	2 T27877	hypothetical prote
13	138	4.5	1385	2 T21706	hypothetical prote
14	137.5	4.5	592	1 LLBY	actin-binding prote

15	137	4.5	724	2 A38749	3-phosphatidylinos
16	135	4.4	3421	1 WZBER6	367k tegument prot
17	134.5	4.4	480	2 JC7552	Shb-like adapter p
18	134.5	4.4	756	2 JC5886	signaling mediator
19	134	4.4	1599	2 T15854	hypothetical prote
20	133	4.4	675	2 S60612	protein-tyrosine k
21	133	4.4	724	2 A38748	3-phosphatidylinos
22	133	4.4	724	2 A38747	phosphatidylinosic
23	132.5	4.3	559	2 T49444	SH3 binding protei
24	132.5	4.3	649	2 T46500	hypothetical prote
25	132	4.3	473	2 S25776	transforming prote
26	130	4.3	670	2 JC5887	signaling mediator
27	129	4.2	751	2 S68957	adhesive plaque pr
28	128	4.2	706	2 A45990	junctional sarcopt
29	128	4.2	1317	2 T03748	apoptosis associat
30	126.5	4.1	1271	1 TVHUR	bcrl (breakpoint cl
31	126	4.1	2440	2 S39162	transcription coac
32	125.5	4.1	596	2 I38228	Shb protein - huma
33	125.5	4.1	2437	2 S53611	MIB1 protein - ra
34	125.5	4.1	2774	2 A43359	microtubule-assoc
35	125	4.1	461	2 A57463	p55PIK - mouse
36	125	4.1	856	2 T16543	hypothetical prote
37	125	4.1	962	2 S58107	hypothetical WD-40
38	125	4.1	1210	2 I39410	AF-4 protein, spli
39	124.5	4.1	585	2 A46209	protein-tyrosine-p
40	124.5	4.1	593	1 JN0805	protein-tyrosine-p
41	124.5	4.1	597	1 A53593	protein-tyrosine-p
42	124.5	4.1	879	2 S64755	hypothetical prote
43	123.5	4.0	733	2 S23468	oocyte-specific pr
44	123.5	4.0	841	2 A43254	protein-tyrosine-p
45	123.5	4.0	1736	2 A47747	tight junction pro

ALIGNMENTS

RESULT 1
B56110
tyrosine phosphoprotein SLP-76 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
C:Accession: B56110
R:Jackman, J.K.; Molto, D.G.; Sun, Q.; Tanemoto, M.; Turk, C.W.; Peltz, G.A.; Koretz
J. Biol. Chem. 270, 7029-7032, 1995
A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated wit
A:Reference number: A56110; MUID:95221345; PMID:7706237
A:Accession: B56110
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <CDS>
A:Cross-references: GB:U20159; NID:9806767; PIDN:AAC52189.1; PID:9806768
C:Superfamily: SH2 homology
C:Keywords: phosphoprotein
F:422-520/Domain: SH2 homology <SH2>

Alignment Scores:

Pred. No.: 3,06e-23 Length: 533
Score: 412.50 Matches: 148
Percent Similarity: 39.88% Conservative: 51
Best Local Similarity: 29.66% Mismatches: 175
Query Match: 13.49% Indels: 125
DB: 2 Gaps: 20

US-09-856-061-1 (1-1721) x B56110 (1-533)
QY 294 TTGGTGATCTGAGATTCAGAACGCTCTCTGCTG-----AAAATAGG 338
||| ||||| ||| |||||
Db 57 PheProlyLeuArgMetProLeuLeuSerLyLeuSerClnsPlleAsnGlu 76
QY 339 TCATGGCCAAAGCTCAGACAGGCGGTGACGGCTTCGACACCTCCG 398
||||| ||||| |||||
Db 77 GluArgSerIlePheThrArgLyProGlnIleProArgPheLeuGluGlnrGlu 96
QY 399 GATCACAGAAAGAACTTGCGCTGGG----- 422

```

Db      97  SerHisGluGlnAspArgLysIleTrpSerSerPheGluAspAspTyrGluSerProAsn 116
      423  -----GTCCAGGTGGAGAAAATGCAACAGTACACAGCTACAGATCT----- 470
Db      117  AspAspAspProAspLysIle-----AspAspGlyAspTyrGluSerProAsnGlu 133
      471  GAGTTCAGCTGCTG-----AAGCATGGCCATCA----- 500
Db      134  GluGluGlnAlaLeuValAspAspAlaAspTyrGluProProSerSerAsnAsnGlu 153
      501  -----ATGAAATTTTACCAGCCAGACCTATCCAG-----GAATCGGAATAC 542
Db      154  GluAlaLeuGlnSerSerIleLeuProProAsnSerPheHisAsnThrAsnSerMetIle 173
      543  CGAGATACACAGCTATTCCAGATATGATGAGGCTCCCTCTGTACTCT-----CCC 596
Db      174  IleAspArgProProThrGlyLysValSerGlnGlnProProValProProLeuArgPro 193
      597  AAGGCTTCTGTC----- 608
Db      194  LysProAlaLeuProProLeuProThrGlyArgAsnHisSerProLeuSerProHis 213
      608  ----- 608
Db      214  ProAsnHisGluGlnProSerArgSerGlyAsnAsnLysThrAlaLysLeuProAlaPro 233
      609  TCCACTGAGAGACAAACAGGATGTAGATGACACAGCTGGAGAGATGGAC----- 662
Db      234  SerIleAspArgSerThrLysProProLeuAspArgSerLeuAlaProLeuAspArgLys 253
      663  -----AAGCTTACCTTCAGAGATGTACAGACCAACGCTTTAAAGCA 704
Db      254  ProPheIleLeuGlyLysLysProProPheSerLysProSerAlaProLeuGlyArg 273
      705  TTCAATATCACAAATAAACAAGACCTCTTGCACCTCTCGGCTTCGTATCACTTC 764
Db      274  GlnHisLeuProLysIleGlnLysProProLeuPro-----ProAlaMetAspArg 290
      765  CCCAAGAGTACCAACCTTACCCCCAGCACACACAGAGAGAGACAGTACTTCCGT 824
Db      291  HisGluArgAsnGluArgLysGlyProValThrThrArg----- 303
      825  CCAAGCCCCACCTTCCAGAGTCCAGAGGGGGCCAGAGAGAGTGCAGAAAGACTTC 884
Db      304  ---LysProSerValProAlaHisGlyArgGlyProAspArgArgGlu----- 318
      885  AGTAGGGTCTTGGAGCAGAGAAGATCTCACACACAGCAAGCCAGAACTCTTCTTC 944
Db      319  -----AsnAspGluAspAspValHisGlnArgProLeuProGlnProSerLeu 334
      945  CCATCATCAACCAAAACACACAGAGAGTCCACTGCCATTTGCCACTCT----- 995
Db      335  ProSerMetSerSerAsnThrPheProSerArgSerValGlnProSerSerLysAsnThr 354
      996  -----TCTTACATGCCGCGAAGACACAGTATACAGCCAGAGACATACAGGTAGC 1046
Db      355  PheProLeuAlaHisMetProGlyAlaPheSer-----GluSerAsnIleGlyPhe 371
      1047  ATGAGCAGACTGT-----CCTGCTCAGAGATGCCAAGCTGCAGCCACAGAGCCCT 1097
Db      372  GlnGlnSerAlaSerLeuProProTyrPheSerGlnGlnLysProGlyAsnAlaGProProLeu 391
      1098  -----CGAATGCTGCCCTATGAAACCAAACTCGGAGAAACCTGCAGCCACCA 1145
Db      392  ArgSerGlnGlyArgAsnLeuProLeuProValProAsn-----ArgProGlnProPro 409
      1146  AAGCCTGATGAGAGAGATGTC-----TGCCAGAAATGAATGAGTACATGGAGATACAGTCGC 1202
Db      410  SerProGlyLysGlnGluIleThrProLeuAspGluGlnIleTyrValSerTyrIleThrArg 429
      1203  CAGGCACTGGAAGATGTTTAATGAAAGACAGAGATGTTTCTTTGGTCCGAGAC 1262

```

```

Db      430  ProGlnAlaGluAlaLeuAlaLeuArgLysIleAsnGlnAspGlyThrPheLeuValArgAsp 449
      1263  TGCCTACAAAATCCAGACAGACATATGTTTGGGTGGTTTATGGAGAAAGTCC 1322
Db      450  SerSerLysLysThrAlaAsnAsnProTyrValIleMetValLeuLysAspLysVal 469
      1323  TACAATGTGAAAATCCGTTTCCCTCGAGACATCAACAGTTTGGCTGGCAGAGACTA 1382
Db      470  TyrAsnIleGlnIleArgTyrGlnGlnGlnSerGlnValTyrLeuLeuGlyThrGlyLeu 489
      1383  CGAGAAATGAGATGTTGTGATTTGGAGAGACATCATTAACATCACTATTTTCCC 1442
Db      490  ArgGlyLysGluAspPheLeuSerValSerAspIleLeuAspTyrPheArgLysMetPro 509
      1443  ATTCTGTAATAGATGGGAAAGACAGGCTCCAGCAGCAAGAAAGTGTACTACACC 1499
Db      510  LeuLeuLeuIleAspGlyLysAsnArgGlySer---ArgTyrGlnCysThrLeuThr 527

RESULT 2
A:56110
tyrosine phosphoprotein SLP-76 - human
C:Species: Homo sapiens (man)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
C:Accession: A56110
R:Jackman, J.K.; Motto, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretz
J. Biol. Chem. 270, 7029-7032, 1995
A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated wit
A:Reference number: A56110; MUID:95221345; PMID:7706237
A:Accession: A56110
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <JAC>
A:Cross-references: CB:020158; NID:9806765; PIDN:AAC50135.1; PID:9806766
A:Genetics:
A:Gene: GDB: LCP2; SLP-76
A:Cross-references: GDB:1230199; OMIM:601603
A:Map position: 5q33.1-5pter
C:Superfamily: SH2 homology
C:Keywords: phosphoprotein
F:422-520/Domain: SH2 homology <SH2>

Alignment Scores:
Pred. No.: 3, 85e-22 Length: 533
Score: 398.00 Matches: 124
Percent Similarity: 42.93% Conservative: 52
Best Local Similarity: 30.24% Mismatches: 186
Query Match: 13.02% Indels: 88
DB: 2 Gaps: 14

US-09-856-061-1 (1-1721) x A56110 (1-533)
      456  GACTACAGATCTCTGATCTCCAGCTGCTGAAGCATGGCCATCAATGAAATTTTACCA 515
      144  AspTyrGlnProProProSerAsnAspGluAlaLeuGlnAsn---SerIleLeuPro 162
      516  GCCAGACCTATCCAGAAATCGGAA-----TACGAGATACACAGCTATTTCCAGATATG 569
      163  AlaLysProPheProAsnSerAsnSerMetTyrIleAspArgProProSerGlyLysThr 182
      570  ATGAGAGCTCC-----CTTCTGTATCTCCAGAGCT 602
      183  ProGlnGlnProProValProProGlnArgProMetAlaAlaLeuProProProAla 202
      602  ----- 602
      203  GlyArgAsnHisSerProLeuProProGlnIleThrAsnHisGluGlnProSerArgSer 222
      603  -----TCTGTCTCCAGTACAGAGCAACAGGGATGTC 635
      223  ArgAsnHisLysThrAlaLysLeuProAlaProSerIleAspArgSerThrLysProPro 242
      636  AGGATGACACAGCTCGAAGAGTGCAC-----AAGCTTACC 671

```


D	b	297	ggtlneuleuserprogllyaenallayrgylalapheluevalalargluserlntlth	317
Oy	1274	ATCCAAAGCGACGACCATTGTGGTGGTG-----TTTTATGGGAACA	1318	
D	b	317	rlysgllyla-----TyrSerLeuSerIleArgAspTrpAspGLIALAArgLYAsPnI	335
Oy	1319	GGTCTCAATGTAATAATCGTTTCCTCGAGAGCATCAACAGTTGGCCCTGGCACAG	1378	
D	b	335	svallyshstYrLyLSleAlrglysleuaprhlgY-----GLTYtTyT	350
Oy	1379	ACTAGCAGGAATGATGCATGTTGATTCCTGTGGAGAAGCATCATTTGAACACTAC	1430	
D	b	350	rllthrThrAlrgAlnPhenAsenValGlnglutleuValGlInHstYc	367
RESULT 4				
T12496				
hypothetical protein DKFZp434I091.1 - human (fragment)				
C:Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 17-Nov-2000				
C:Accession: T12496; T46492				
R:Poultk, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.				
submitted to the Protein Sequence Database, June 1999				
A:Accession: T12496				
A>Status: preliminary				
A:Molecule type: mRNA				
A:Residues: 1-572 <POD>				
A:Cross-References: EMBL:AL080174				
A:Experimental source: adult testis; clone DKFZp434I091				
R:Poultk, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.				
submitted to the Protein Sequence Database, January 2000				
A:Reference number: Z23029				
A:Accession: T46492				
A>Status: preliminary				
A:Molecule type: mRNA				
A:Residues: 456-572 <AAA>				
A:Cross-References: EMBL:AL137328				
A:Experimental source: adult testis; clone DKFZp434A2372				
C:Genetics:				
A>Note: DKFZp434I091.1; DKFZp434A2372.1				
Alignment Scores:				
Pred. NO.:	0.00276	Length:	572	
Score:	149.50	Matches:	106	
Percent Similarity:	33.59%	Conservative:	66	
Best Local Similarity:	20.70%	Mismatches:	200	
Query Match:	4.89%	Indels:	140	
Db:	2	Gaps:	25	
US-09-856-061-1 (1-1721) x T12496 (1-572)				
Oy	240	CAGGTGCCAGACGACATGACGACCCAGGGCATTAAGACACAGAA---GAAGCATTC	296	
D	b	131	glnleuproproglthrArProAlaglyAsnLysaspgrugluglutyrgllylph	150
Oy	297	GGTGATCTGAGATTCGAAACGCTCTCTCGCTGGAANAATAGATCATGCCAACGCCACAG	356	
D	b	151	-----serlleieserAnglurProleuAlarhtYrGln	162
Oy	357	AGTCCAAAGGGGGGTGTGCAGCGGTTCGGAACCACTTCGGATTCACAGAGAAGCACTTG	416	
D	b	163	gluAmnArfyaSnThrcyspheser-----ProasPhISLysglnProgln	177
Oy	417	GCTGGGGTCCAGAGTGGAGAAAA-----TGCACAGACTAACAMCGACTCAAGAAATGCT	470	
D	b	178	proglylThralaelgylwetaArgAlaserValserSerLysglYserlleSerThrPro	197
Oy	471	GAGTTCACAGTGTGCTGAAGCATGCGCATCAATGAANAATTTTACCAGGC-----AGACCT	524	
D	b	198	AasnSerProval-----LysmethThrleuileProSerProAsnSerPro	212
Oy	525	ATCCAG-----GAATCGGAATTCGACGATACA	551	

[illegible]

QY 1052 GCACGTCTGCTGCTCAGAGATGCCAAGCTGACGACGACCCCTGCAATGCTGCCCTA 1111
 Db 1441 -----ProaspiliserProgl 1446
 QY 1112 TGAACAACAACTCGGAAACCTGACCCCAAAAGCTGTAGAGATGCTGGCA 1171
 Db 1446 uThraArgAlaAsnValLysPheValGlnAspThr-----Se 1458
 QY 1172 GAATGAATGCTACATTTGGAGATACAGTCCG---CAGCAGTGGAGATGTTAATGAA 1228
 Db 1458 rlystyrttpttlyrlyspProaspiliserArgLglnAlaIle-----AlaLeuLeuLys 1476
 QY 1229 AGAGAACAAAGATGCTACTTTTGGTCCGACAGCTGCTTACAAATGCCAAGCAGAAC 1288
 Db 1476 sasprtgglupProglYAlaPheIleIleArgspSerHisSer----- 1490
 QY 1289 AATATGTTTGGTGTGTTTATGGAACAAGCTCTACAAATGTAATCCGTTTCCTGA 1348
 Db 1491 -----PheArgGly-----AlaTyrGlyLeuAlaMetLysValAlaLase 1503
 QY 1349 GAGCAATCAACAGTTTGCCCTGGCAGACGACTACGAGAAATGAGATGTTGATCTGT 1408
 Db 1503 rProProPro-----ThValMetGlnGlnAsnLysLysGlyAspIleTh 1518
 QY 1409 GGAACACATCATTTGACACTACACATATTTTCCATTCTGTAATAGATGGAAAGACAA 1468
 Db 1518 rAsnGlnLeuValArgHisPhe-----LeuIleGlnThrSerProAr 1532
 QY 1469 GGCTGCACGACGAGAAACAGTGC----- 1490
 Db 1532 gGlyValLysLeuLysGlyrlyspProaspiliserProaspiliserPheGlyLysLeuVal 1552
 QY 1491 -TACTCACCACGACCTGCTCTGCGCAGGCTCTTCTCACTAGTACTCAGCAGC 1549
 Db 1552 lTyrGlnHisSerIleMetProLeuAlaLeu-ProCysLysLeuValIleProaspiliser 1572
 QY 1550 ACTTCAGTAAAGACCCGACGACGAGATAGCCGACATGCGCTGGCTGTCTCA 1609
 Db 1572 sProMetGlnGlnLysLysAspAlaLase----- 1582
 QY 1610 TTCCTGCTCAATGATGATCTCTTCTTCATCTGCAATTTATCTGCAATATATTTT 1669
 Db 1583 --rThrHisAsnSerAlaThrAspLeuLeuLysGlnGlyAlaAlaCysAsnValLeuPhe 1602
 QY 1670 CTGTGCTTCAAGGATGATTTTGTGACT 1698
 Db 1602 leAsnSerValGlnMetGlnSerLeuThr 1611
 RESULT 9
 A54970
 tensin, cardiac muscle - chicken
 C:Species: Gallus gallus (chicken)
 C:Date: 11-Nov-1994 #sequence, revision 11-Nov-1994 #text, change 21-Jul-2000
 C:Accession: A54970; S38330; S21544
 R:Id, S.H.; An, Q.; Bao, S.; Wong, W.K.; Liu, Y.; Janney, P.A.; Hartwig, J.H.; Chen, L.H.
 J. Biol. Chem. 269, 22310-22319, 1994
 A:Title: Molecular cloning of chick cardiac muscle tensin. Full-length cDNA sequence, ex
 A:Reference number: A54970; MUID:9430987; PMID:8071358
 A:Accession: A54970
 A:Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra
 A:Molecule type: mRNA
 A:Residues: 1-1744 <LOA>
 A:Cross-references: GB:M96625
 R:van de Werken, R.; Gennari, M.; Tavelle, S.; Bel, P.; Molina, F.; Lin, S.; Canceda, F
 Eur. J. Biochem. 217, 781-790, 1993
 A:Title: Modulation of tensin and vimentin expression in chick embryo developing cartila
 A:Reference number: S38330; MUID:94039118; PMID:8223621
 A:Accession: S38330
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1469-1744 <VAP>
 A:Cross-references: EMBL:X66286; NID:963802; PIDN:CAA46992.1; PID:g63803
 C:superfamily: SH2 homology

C:Keywords: cardiac muscle; heart
 F:1472-1581/Domain: SH2 homology <SH2>
 Alignment Scores:
 Pred. No.: 0.0177 Length: 1744
 Score: 140.00 Matches: 93
 Percent Similarity: 34.11% Conservative: 54
 Best Local Similarity: 21.58% Mismatches: 138
 Query Match: 4.58% Indels: 146
 DB: 2 Gaps: 21
 US-09-856-061-1 (1-1721) x A54970 (1-1744)
 QY 527 CCAGAAATCGGAATACGACGATACACGCTATTTCCAGATATGATGAGCGCTCCCTTCT 586
 Db 1299 ProGlyLeuGly-----ArgHisThrValSerProHisAlaProProGlySerProSer 1316
 QY 587 GTTACCTCCCA----- 610
 Db 1317 LeuAlaArgHisGlnMetAlaAlaValProProGlySerProMetTyrGlyTyrSer--Se 1336
 QY 611 CACTGACGACAAACAGGATGTAGAGATGACACAGCTGGAAAGTGGCAAGCCTAC 670
 Db 1336 rProGlnGlnArg-----ArgProth 1343
 QY 671 CTTCAGAGATGTCAGAAACCCCTTTAAAGATTCAATACAAATAAACAAGAC 730
 Db 1343 rLeuSer-----ArgLysSerAlaSerGlyTyrGln----- 1354
 QY 731 TCCCTTCCACCTCTGCGCTGTATCACTCTCCCAAGAACAGACCAACC----- 782
 Db 1355 -----ProProSerThrProSerPheProValSerProAlaTyrTyrProGlyLysrse 1372
 QY 783 -TTACCCCAAGCAGCAGCAGAGAGACAGTGCATCTTC-----GCTCCAA 829
 Db 1372 rThrProHisSerSerSerProaspiliserAlaAlaTyrArgGlnGlySerProThrProgl 1392
 QY 830 GCCCACCCTTTCAGAAATCCAGAG-----GGGCCAGGACAGAGGTGCMAAAG 880
 Db 1392 nProAlaLeuProGlnLysArgArgMetSerAlaGlyArgLysSerHisSerLeuPro 1412
 QY 881 CTTCAGTAAAGCTCTGAGCAGACAGAAAGATCTCACACCAACAAAGCCAGATCTTC 940
 Db 1412 nTyrAlaThrValAsnGly-----LysLase 1421
 QY 941 TTGCCCA---TCATCAACCAAAACACACAGAAAGTCCACCTGCCATGCCCTTC 997
 Db 1421 rSerProLeuSerSerGlyMetSerSerProSerSerGlySerAlaValAlaPheSerHis 1441
 QY 998 CTACATGCCA-----GGAACGACAGATATCAACACGACAGACCATACAGTATGATGCA 1051
 Db 1441 sThrLeuProAspPheSerLysPheSerMet----- 1451
 QY 1052 GCACGTCTGCTGCTCAGAGATGCCAAGCTGACCCCAAAAGCTGTAGAGATGCTGGCA 1111
 Db 1452 -----ProaspiliserProgl 1457
 QY 1112 TGAACAACAACTCGGAAACCTGACCCCAAAAGCTGTAGAGATGCTGGCA 1171
 Db 1457 uThraArgAlaAsnValLysPheValGlnAspThr-----Se 1469
 QY 1172 GAATGAATGCTACATTTGGAGATACAGTCCG---CAGCAGTGGAGATGTTAATGAA 1228
 Db 1469 rlystyrttpttlyrlyspProaspiliserArgLglnAlaIle-----AlaLeuLeuLys 1487
 QY 1229 AGAGAACAAAGATGCTACTTTTGGTCCGACAGCTGCTTACAAATGCCAAGCAGAAC 1288
 Db 1487 sasprtgglupProglYAlaPheIleIleArgspSerHisSer----- 1501
 QY 1289 AATATGTTTGGTGTGTTTATGGAACAAGCTCTACAAATGTAATCCGTTTCCTGA 1348
 Db 1502 -----PheArgGly-----AlaTyrGlyLeuAlaMetLysValAlaLase 1514


```

Db 329 -----AlaProPProSerProLeuSerMetSerGluValProIleAl 342
OY 1010 AAGACAGATATACACACAGACACATACAGTACGATCAGACACTGCTCTCAGAG 1069
Db 342 AAlaHisGluValGluIleGluAsp ----- 350
OY 1070 ATGCAGAGTGCAGACCCAGCCCTGCATGCTGCCCTATGAAACCAACTCGGA 1129
Db 351 ----GluValValGlnSerGlySerProIleAspIleValThGluValGlnSerAsnAs 369
OY 1130 GAACCTGACCCCAAAAGCCTGATGAGAGAGATGCTGCAGATGAATGATACATGG 1189
Db 369 pGlyProProMetProPro-----MetLeuLeuG 380
OY 1190 AGAATACAGTCCGACG----- 1205
Db 380 nGlnGlnSerSerGlnAlaSerSerSerThryAlProSerAlaSerGlnGlyValThryA 400
OY 1206 -----GCAGTGAAGATGCTTAAATGAAGAGAACAGAGATGATCTTTTGGT 1255
Db 400 lProSerThryAlaLeuAspAlaThryAlGlySerGluValGlnGlyLysAspLeuPr 420
OY 1256 CCGAGACTGCTCTACAAATCCAGGCA-----GAACCATTA 1291
Db 420 oSerGluHisPheProLysGluLysAlaArgHisValSerThryArgLysLeuHisProTy 440
OY 1292 TGTGTGTGTGTGTGTGTATGAGCAACAG-----GTCATCAATGGAATACCCGTTTCT 1345
Db 440 rVal-----SerAspLysSerHisIleIleThryAspValAlaGlyAspGlnValLe 455
OY 1346 C-----GAGACGAATCAACAGATTTGCGCCGCGCC-----AC 1375
Db 455 ulleHisProGluAspMetGlnGlnGlnGlnAlaIleAlaGlnGlnArgValLe 475
OY 1376 AGGACTACAGCAAAATGAGATGTT-----GATTCGTGAGAGA 1414
Db 475 uThrLeuGlnGlyGlyAsnPheIleGluAspArgIleGlyArgSerAspLeuMetGluAs 495
OY 1415 C-----ATCATTTGAACAC----- 1427
Db 495 pGlnGluMetLysMetLysSerIleGlnProGlnGlnIleValGlnGluAsnGln 515
OY 1428 -----TACACATATTTCCCATCTCTGCTAATAGATGGCAAGCAAGGCT-- 1472
Db 515 yMetLysValValGlyGluValAspProAsnSerValIleGlnGlnGlnGlnValGln 535
OY 1473 -----GCACGCAAGAAACAGTGTCTACTCACCAGCCACTGCGCTCGCCAGAGCTCT 1525
Db 535 uIleGlnAlaIleGlnGlnGlnAla-----GluGlnIleThryA 547
OY 1526 TCTCAGTACAGTACCTCCAGCCAGCATTCATGAGTAAGAACCCAGCCAGATATCCCC 1583
Db 547 lMetAspGlnHisSerLysGlnGlnGlnMetGluAspGlnGluAspAspLeuPro 566

RESULT 14
LBLY
actin-binding protein ABPI - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YCR088w
C:Species: Saccharomyces cerevisiae
C:Date: 31-Mar-1991 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
C:Accession: S19503; S19767; S07608
R:Dusterhoft, A.; Erdmann, D.; Heemann, J.; Philippson, P.; Schweitzer, B.; Spiegelberg
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19503
A:Molecule type: DNA
A:Residues: 1-476 <DUS>
A:Cross-references: EMBL:X59720; GSPDB:GN00003; MIPS:YCR088w
R:Fromail, L.; Grisanti, P.
submitted to the Protein Sequence Database, March 1992
A:Reference number: S19504
A:Accession: S19767

```

```

A:Molecule type: DNA
A:Residues: 14-592 <FRO>
A:Cross-references: EMBL:X59720; GSPDB:GN00003; MIPS:YCR088w
R:Drubin, D.G.; Mulholland, J.; Zhu, Z.; Botstein, D.
A:Title: Homology of a yeast actin-binding protein to signal transduction proteins an
A:Reference number: S07608; MUID:90136906; PMID:2405279
A:Molecule type: DNA
A:Residues: 1-57,'S',59-311,'I',313-592 <DRU>
A:Cross-references: EMBL:X51780; NID:g3321; PIDN:CAA36075.1; PID:g3322
C:Genetics:
A:Gene: SGD:ABPI; MIPS:YCR088w
A:Cross-references: SGD:S0000684; MIPS:YCR088w
A:Map position: 3R
C:Superfamily: actin-binding protein ABPI, SH3 homology
C:Keywords: actin binding; duplication
F:158-180,378-401/Region: duplication
F:539-588/Domain: SH3 homology <SH3>

Alignment Scores:
Score: No.: 0.0226 Length: 592
Percent Similarity: 137.50 Matches: 84
Best Local Similarity: 36.88% Conservative: 58
Query Match: 21.82% Mismatches: 135
DB: 4.50% Indels: 108
Gaps: 20

US-09-856-061-1 (1-1721) x LBLY (1-592)
OY 221 GAGATCCCTCTCTTTGCGCCAGGTCACAGACCATGACAGCCAGCAATTAAGAC 280
Db 244 GlusProAlaGlyLeuVal-GlnLysPro-----ThryAlaIleGlySerLysIleAs 260
OY 281 AACGAAGAGATTCGCTGATCTGATCTCAGATTCACAGACGTCCTGCTGAAATAGTC 340
Db 260 pProSerSerAspIleAlaAsnLeuLys-----AsnGlnSerLysLeuLysArgAspSe 278
OY 341 ATGCCAAGCTCAGCAGCTGCCAAGAGCGCGGTGCGAGCGGTCTGGAACCATTCGGA 400
Db 278 r-----GluPheAsnSerPheLeuGlyThrThryPro----- 289
OY 401 TCACAGAAGCAATTTGGCTGGGCTGCCAGGTGGAGAAATGCAACAGTAAACAGACTA 460
Db 290 -----ProSerMetThrGlnSerSerLeuLysAsnAsp-- 300
OY 461 CGAAGATCTGAGTTCAGCTGCTGAAGCATG-----CCATCAATGAA 505
Db 301 -----AspAspLysValIleLysGlyPheArgAsnGlnLysSerProAlaGln 317
OY 506 AATTTACAGCCAGCCTATCCAGAGATGGAATAGCAGATACAGCTATTTCAGCA 565
Db 317 uTPAlaGluArgLysAlaLysGlnAsnSerGlyAsnAlaGluThryLys----- 333
OY 566 TATGATGAGAGCTCCCTCTCTGTTACCTCCCAAGGCTTCTGCTCCAGTGAAGACAAC 625
Db 334 ----AlaGluAla-----ProLysProGluValProGluAsp----- 344
OY 626 CAGGATGTGAGATGATACACAGCTGAGACAGTGGCAAGACCTTACCTCAAGAGTGCAG 685
Db 345 -----GluProGlnGlyGlnProAspValLysAspLeuLys 356
OY 686 AACCAAGCTTTAAGATTCATTAATACCAAAATTAACAGACTCCCTTG----- 737
Db 356 sSer-----LysPheGlnGlyLeuAlaIleSerGlnGlnGlnGlnGlnGlnGlnGln 735
OY 738 -----CACTCTCTGCG-----CTGCTATACACTCCCAAGAGTAC-----CA 778
Db 375 sPheAlaProProLysLysSerGlnProThrIleIleSerProLysProPheSerLys 395
OY 779 ACCCTTACCCCAAGCAGCAAGAGAGAGAGAGTGCATCTTCGCTCCCAAGCCACTT 838
Db 395 sProGlnGlnProValLysAlaGlnGlnAlaGlu-----GlnProLysThryAspTy 412

```


Db 379 ysleuilelysilphenhisargaspGlytyrGlypheserasprou----- 396

QY 1390 ATGAGATGTTGATCTCTGAGAGACATCATGAACACTAC 1430

Db 397 -----ThrpheasnservalValGluLeuileasnHisTyr 408

Search completed: April 21, 2003, 12:32:00
Job time : 89.7435 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_n2p model

Run on: April 21, 2003, 12:15:11 ; Search time 24.7582 Seconds
(without alignments)
5766.221 Million cell updates/sec

Title: US-09-856-061-1
Perfect score: 3057
Sequence: 1 acgagcgccaactgcacag.....taaaaaaaaaaaaaaaaaa 1721

Scoring table:
BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 112892 segs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame_n2p.model -DEV=xlp
-Q/cgn2.1/USPTO-spool/US09856061/runat-21042003.113018.3735/app.query.fasta-1.3150
-DB=SwissProt-40 -OFMT=fasta -SUFFIX=n2p.rsp -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09856061.ecgn.1.1.65.etrunal-21042003.113018.3735 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG-SCORES=0 -WAIT -LONLOG -DEV-TIMEOUT=120
-MARN-TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412.5	13.5	533	1	LCP2_MOUSE
2	398	13.0	533	1	LCP2_HUMAN
3	148	4.8	722	1	P85B_RAT
4	146.5	4.8	217	1	GRAP_HUMAN
5	146.5	4.8	724	1	P85B_BOVIN
6	146	4.8	545	1	FGR_FSVGR
7	146	4.8	728	1	P85B_HUMAN
8	140.5	4.6	540	1	SCK_HUMAN
9	140	4.6	1744	1	TENS_CHICK
10	139.5	4.6	1603	1	PSC_DROME
11	137.5	4.5	592	1	ABP1_YEAST
12	137.5	4.5	722	1	P85A_MOUSE
13	137	4.5	724	1	P85A_BOVIN
14	136.5	4.5	561	1	3BP2_HUMAN
15	135	4.4	3421	1	TG01_HSVB
16	133	4.4	675	1	BMX_HUMAN
17	133	4.4	724	1	P85A_HUMAN
18	133	4.4	724	1	P85A_MOUSE

19	133	4.4	724	1	P85A_RAT	063787	rattus norv
20	132.5	4.3	271	1	CRK_DROME	09xym0	drosophila
21	132.5	4.3	559	1	3BP2_MOUSE	006649	mus musculu
22	132	4.3	583	1	SHC_HUMAN	P29353	homo sapien
23	129	4.2	461	1	P55G_HUMAN	092569	homo sapien
24	129	4.2	751	1	FPI_MYTGA	027409	mytilus gal
25	128.5	4.2	901	1	T298_HUMAN	015016	homo sapien
26	128	4.2	705	1	TRDN_RABIT	028820	cricetolagus
27	127	4.2	705	1	P55G_BOVIN	064404	bos taurus
28	127	4.2	1362	1	BRD4_HUMAN	060885	homo sapien
29	127	4.2	1386	1	ZAP3_MOUSE	093017	mus musculu
30	126.5	4.1	1271	1	BCR_HUMAN	P11274	homo sapien
31	125.5	4.1	2774	1	MAPA_RAT	P34926	rattus norv
32	125	4.1	962	1	P55G_MOUSE	064143	mus musculu
33	125	4.1	962	1	Y4E_SCHPO	009731	schizosacch
34	125	4.1	1210	1	AR4_HUMAN	P51825	homo sapien
35	124.5	4.1	583	1	PTNB_MOUSE	P55235	mus musculu
36	124.5	4.1	593	1	PTNB_RAT	006123	homo sapien
37	124.5	4.1	593	1	PTNB_RAT	P41499	rattus norv
38	124.5	4.1	1490	1	CRK7_HUMAN	09nyv4	homo sapien
39	123.5	4.0	845	1	CSM_DROME	P29349	drosophila
40	123.5	4.0	1736	1	Z01_HUMAN	007157	homo sapien
41	122.5	4.0	875	1	FPI_MYTED	025460	mytilus edu
42	121.5	4.0	1147	1	K1N2_YEAST	P13186	saccharomyc
43	121	4.0	1217	1	AR4_MOUSE	088573	mus musculu
44	121	4.0	2442	1	CBP_HUMAN	092793	homo sapien
45	120.5	3.9	806	1	MK07_MOUSE	09wv58	mus musculu

ALIGNMENTS

RESULT 1

LCP2_MOUSE
ID LCP2_MOUSE STANDARD: PRT: 533 AA.

AC 060787;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte
protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
GN LCP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-T-cell lymphoma;
RX MEDLINE=95221345; PubMed=7706237;
RA Jackson J.K., Motto D.G., Sun Q., Tanemoto M., Turck C.W., Peltz G.A.,
Koretzky G.A., Finkel P.R.;
RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein
associated with Grb2 in T cells";
RT J. Biol. Chem. 270:7029-7032(1995).
CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND FYB.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND
PERIPHERAL BLOOD LEUKOCYTES.
CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION (BY
SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC
CC EMBL: U20159; AAC52189.1; --
CC HSSP: P12931; ISHD.

DR MGD; MG1:1321402; LCP2.
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00454; SAM; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS0001; SH2; 1.
 DR SH2 domain; Phosphorylation.
 FT DOMAIN 422 530
 SEQUENCE 533 AA: 60228 MW: 6DD877782C2660C3 CRC64;

Alignment Scores:
 Pred. No.: 7,11e-21 Length: 533
 Score: 412.50 Matches: 148
 Percent Similarity: 39.88% Conservative: 51
 Best Local Similarity: 29.65% Mismatches: 175
 Query Match: 13.49% Indels: 125
 DB: 1 Gaps: 20

US-09-856-061-1 (1-1721) x LCP2_MOUSE (1-533)

OY 294 TTCGGTGTGATGAGATTCAGAACGCTCTCTGCTG-----AAAAATAG 338
 DB 57 PHEPRLYSLLEAARMETCPLEULEUSERLYSLSESGINASPILLEASNLCSNGLU 76
 OY 339 TCATGGCCAAAGCCCTGAGCAGCCAAAGGCGGTGTCGAGCGTTCGAAACCACTTCG 398
 DB 77 GIATGATGSERLIEPHEHTRARGLYSPROGINLEPROARGPHELEUGLUHTRGLU 96
 OY 399 GATCAGACAGAACTTGGCTGG----- 422
 DB 97 SERHISGLUGLUAASPAPGLYLITRPSERSEPHGLUASPAPTYRGLUSERPROASN 116
 OY 433 -----GTCCGAGGAGCAAAATGCAACAGTACACAGCTAGCAAGATCC----- 470
 DB 117 ASPAPAPAPROAPGLYGLU-----ASPAPGLYASPTRYGLUSERPROASNGLU 133
 OY 471 GAGTTCACGCTGCTG-----AAGGATGGCCATCA----- 500
 DB 134 GINGLUGLUALALEUVALASPAPALALAAPRYRGLUPROPROSEFASASNGLU 153
 OY 501 -----ATGAATAATTTACAGCCAGCACTATCCAG-----GAATCGAATAC 542
 DB 154 GLUALALEUGLINSERSERLIELEUPROPROASNSEPHENIASNTHASNSERMETLYR 173
 OY 543 GCAGATACAGCCTATTTCAGGATATGATGAGGCTCCCTTGTACT-----CCC 596
 DB 174 ILEAPATGPROPTHGLYLVALSERGININPROPROVALPROPROLEUARGPRO 193
 OY 597 AAGGCTCTGTC----- 608
 DB 194 LYSPPALALEUPROPROLEUPROTHRGLYARGASNHSERPROLEUSERPROPROHS 213
 OY 608 ----- 608
 DB 214 PROASNHSGLUGLUPROSERARGSEGLYASNSLSTHRLALYSLLEUPROALAPRO 233
 OY 609 TCACCTGAGACAAACCCAGGATGTGAGATGACACAGCTGGAAGAATGAGC----- 662
 DB 234 SERILEAPATGSERLTHLYSPROPROLEUASPARGSERILEUALPROLEUASPARGLU 253
 OY 663 -----AAGCCTACCTTCAGAGATGTGAGAACCCAAACGCTTTAAAGA 704
 DB 254 PROPHLEULEUGLYLYSLYSPROPROPHESERASPLYSERPROVALAPROLEUGLYARG 273
 OY 705 TTCAATATACAAATAAACAAGATCTTGGCACCCTCTGCGCTGATACACTCTC 764
 DB 274 GLUHISLEUPROLYSILEGLINLSPROPROLEUPRO-----PROALAMEASPARG 290
 OY 765 CCCAAGAAGTACCAACCTTACCCAGACACACAGAGAGAGAGACTGATCTGCTGCT 824
 DB 291 HISGLUARGASNGLUARGLEUGLYPROVALTHTRARG----- 303

OY 825 CCAAGCCACCTTTCAGAACTCCAGAGGGGCCAGGAGAGATGCAAAAGACTTC 884
 DB 304 ---LYSPROSERVALPROARGHISGLYARGLYPROAPARGGLU----- 318
 OY 885 AGTAGGTCCTTGAGCAGAGAAAGATCTCCACCCAGACAAACCGAATCTTCTTC 944
 DB 319 -----ASNAPGLUASPAPVALHISGLINARGPROLEUPROGLINPROSERLEU 334
 OY 945 CCATCATCAACCAACAGACAGAGATCCACCTGCAATTCAGCCTC----- 995
 DB 335 PROSERMETSERSEFASNTHRPHROSERARGSERVALGINPROSERSELYASNTHR 354
 OY 996 -----TCTACATCCAGAGAAAGCAGATATACAGCCAGACCATACAGTAGC 1046
 DB 355 PHEPROLEUALAHISMETPROGLYALAPHESE-----GLUSERASNILEGLYPHE 371
 OY 1047 ATGACGACTGT-----CCTGCTCAGAGATGCCAAGCTGCAGCCAGCACGCCCT 1097
 DB 372 GINGLINSERIALSERLEUPROPROTYRPHSESERGINLYPROGLYASNARGPROLEU 391
 OY 1098 -----CGAATGCTCCCTATGAAACAAACAGCTGAGAAACCTGACCCCA 1145
 DB 392 ARGSERLUGLYARGASNLEUPROLEUPROVALPROASN-----ARGPROGLINPRO 409
 OY 1146 AAGCCTGATGAGAGAGATGTC---TGCAGAAATGAAATGATTCATTGCAATACAGTCC 1202
 DB 410 SERPROGLYGLUGLUGLUTHRPROLEUASPGLUUTRPTRYVALSERTRYLLETHARG 429
 OY 1203 CAGGCAATGAGATGTGTTATGATGAAGACAGAAAGATGTTACTTTTGTGTCGAGAC 1262
 DB 430 PROGLUALAGLUALALEUARGLYSLLEASNINSPGLYTHRPHLEUVALARGASP 449
 OY 1263 TGCTCTCAAAATCCAAAGCCAGAACATATGTTTGTGTTTATGGAACAAGCTC 1322
 DB 450 SERSERLYSLTHRLAASASNAPROTYRVALLEUMETVALLEUTRYLYSASPLYVAL 469
 OY 1323 TACAATGGAATAATCCGTTCTCGAGAGCAATACAGTTGGCCGCGACAGCACTA 1382
 DB 470 TYRASNILEGINLIEAGTYRGLINGLUGLUSERGINVALYTRLEUAGLYTHGLYEU 489
 OY 1383 CGAGAAATGAGATGTTGATCTGTGAGACATCATTCGAACACTACATATTTTCCC 1442
 DB 490 ARGGLYSLGILUAAPHELEUSERVALSERAPLIEILEASPYRPHARGLYSMEPRO 509
 OY 1443 ATCTGCTAATAGATGGGAAGAACAGAGCTGCAGCAGAAACGTCTACTTACC 1499
 DB 510 LEULEULEULEASPLYLYSASNARGLYSER---ARGTYRGLINCYSTRHLEUTHR 527

RESULT 2
 LCP2_HUMAN
 ID LCP2_HUMAN STANDARD: PRT: 533 AA.
 AC 013094:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte
 DE protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
 GN LCP2.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Leukemia;
 RX MEDLINE=95221345; PubMed=7706237;
 RA JACKMAN J.K., MOTO D.G., SUN Q., TANEMOTO M., TURK C.W., PELTZ G.A.,
 RA KORECZY G.A., FINDERLIN P.R.;
 RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein
 RT associated with Gb2 in T cells.";
 RL J. Biol. Chem. 270:7029-7032(1995).
 RJ [2]

RP SEQUENCE FROM N.A.
 RC Tissue-Prostate;
 RA Strausberg R.1
 RL Submitted (Oct-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
 CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND Fyb.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND
 CC PERIPHERAL BLOOD LEUKOCYTES. HIGHLY EXPRESSED ALSO IN T CELL AND
 CC MONOCYTIC CELL LINES, EXPRESSED AT LOWER LEVEL IN B CELL LINES.
 CC NOT DETECTED IN FIBROBLAST OR NEUROBLASTOMA CELL LINES.
 CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION BY Zap-70.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- DATABASE: NAME-PROT; NOTE-PROT 1:1-5(2000);
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1118450040_g.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: U20158; AAC50135.1; -
 DR EMBL: BC016618; AAH16618.1; -
 DR HSSP: P12931; 1SHD.
 DR Genew; HGNC:6529; LCP2.
 DR MIM: 601603; -
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART; SM00454; SAM; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS50001; SH2; 1.
 KW SH2 domain; Phosphorylation.
 FT DOMAIN 133 136 POLY-GLU.
 FT DOMAIN 198 201 POLY-PRO.
 FT DOMAIN 422 530 SH2.
 SQ SEQUENCE 533 AA; 60188 MW; C5D22P31D36200C8 CRC64;
 Alignment Scores:
 Pred. No.: 6,996-20 Length: 533
 Score: 398.00 Matches: 124
 Percent Similarity: 42.93% Conservative: 52
 Best Local Similarity: 30.24% Mismatches: 146
 Query Match: 13.02% Indels: 88
 DB: Gaps: 14
 US-09-856-061-1 (1-1721) x LCP2_HUMAN (1-533)
 QY 456 GACTACGAGAGATCCGAGTTCAGCTGAGAGGATGGCCATCAATGAAATTTTACCA 515
 DB 144 AsptylgluProProProSerAsnAspGluAlaLeuGlnAsn---SerIleLeuPro 162
 QY 516 GCGAGACCATCCAGAAATCGGAA-----TAGCCAGATACACCGCATTTCCAGATATG 569
 DB 163 AlatyProPheProAsnSerAsnSerMetTyrIleAspArgProProSerGlyLysThr 182
 QY 570 ATGAGAGCTCC-----CTTCTGTACTCCCAAGGCT 602
 DB 183 ProGlnGlnProProValProProGlnArgProMetAlaAlaLeuProProProAla 202
 QY 602 ----- 602
 DB 203 GlyArgAsnHisSerProLeuProProGlnThrAsnHisGluGluProSerArgSer 222
 QY 603 -----TCGCTCCACGAGAGACAAACAGGATGTG 635
 DB 223 ArgAsnHisLysThrAlaLysLeuProValProSerIleAspArgSerThrLysProPro 242
 QY 636 AGGATGACACAGCTGGAGAAAGTGGAC-----AAGCTTACC 671

DB 243 LeuAspArgSerLeuAlaProPheAspArgGluProPheThrLeuGlyLysLeuProPro 262
 QY 672 TTCAGAGAT-----GTGAGAAAGCCACAGCCTTTAAAGATTTAAATACACAAATA 722
 DB 263 PheAsnAspLysProSerIleThrAlaGluArgSerLeuGlyGlnIleLeuProLysIle 282
 QY 723 AACAGATCCCTTTCACATCCCTCGCTGGCTGTATCACTCTCCCAAGATACCAACC 782
 DB 283 GlnLysProProLeuProPro-----ThrThrGlnArgHisGluArg 296
 QY 783 TTACCCCAACACACACAG 842
 DB 297 SerSerProLeuPro-----GlyLysLysProProValPro 308
 QY 843 GAAGTCCAGAGGGGGCCAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 902
 DB 309 LysHisGlyTyrGlyProAspArgArgGlu-----AsnAsp 320
 QY 903 GAGAGAAATCTCACACACAG 944
 DB 321 GluAspAspValHisGlnArgProLeuProGlnProAlaLeuLeuProMetSerSerAsn 340
 QY 945 CCATCATCAAAACCAAAACACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1004
 DB 341 ThrPheProSerArgSerThrLysProSerProMetAsnProLeuProSerSerHisMet 360
 QY 1005 CCAGAGAAAGCAGATATACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1064
 DB 361 ProGlyAlaPheSer---GluSerAsnSerSerPheProGlnSerAlaSerLeuProPro 379
 QY 1065 CAGAGATGCCAGCTGCAG 1109
 DB 380 TyrPheSerGlnGlyProSerAsnArgProProIleArgAlaGluGlyArgAsnPhePro 399
 QY 1110 TATGAAACCAAACTCGAG 1169
 DB 400 LeuProLeuProAsnLysProArgPro---ProSerProAlaGluGluLysAsnSerLeu 418
 QY 1170 CAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1229
 DB 419 AsnGluGluTyrTyrValSerTyrIleThrArgProGluAlaLeuAlaLeuArgLys 438
 QY 1230 GAGAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1289
 DB 439 IleAsnGlnAspLysThrPheLeuValArgAspSerSerLysLysThrThrAsnPro 458
 QY 1290 TATGTTTGGTGTGTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1349
 DB 459 TyrValLeuMetValLeuTyrLysAspLysValTyrAsnIleGlnIleArgTyrGlnLys 478
 QY 1350 AGCAATCAACAGTTTCCCTGGGACAGAGATACAGAGAGAGAGAGAGAGAGAGAGAGAG 1409
 DB 479 GluSerGlnValTyrLeuLeuGlyThrGlyLeuArgGlyLysLysAspPheLeuSerVal 498
 QY 1410 GAAGACATCATTCAGATACATATTTTCCATCTCTCTATATAGATGGAAGAGCAAG 1469
 DB 499 SerAspIleIleAspTyrThrArgLysMetProLeuLeuLeuIleAspLysAsnArg 518
 QY 1470 GCTGACAGCAGAAACAGTCTACTACC 1499
 DB 519 GlySer---ArgTyrGlnCysThrLeuThr 527
 RESULT 3
 P85B-RAT
 ID P85B-RAT STANDARD; PRT: 722 AA.
 AC 063788;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 DE p85-beta subunit) (PcGlns-3-kinase p85-beta).
 GN PIK3R2.

Db	55	gltpr-metprgilypheasngluarqthrtarglnarqglyasphernproglythr	74
QY	471	-----GAGTTC-----CAGCTGCTAAGCATGGCCA-----TCATGAA	505
Db	74	rvalgnupheleuglpyrovalalalelaalaaryproglpyroargproargglyproar	94
QY	506	AATTTTACAGCCAGCAACCTATCCAG---GAATCGAATTAAGCAGATPCCA---CGCATATT	559
Db	94	groleupproalaargproleuabarglyprosergluserclylnsthrleualaaserle	114
QY	560	CCAGGATATGATGAGAGCTCCCTCTGTACTCC-----	596
Db	114	uallaglnpheserproprogluseralraproproilleuvalylsleuileglnal	134
QY	597	-----AAGGCTTCTCTCCATCGAGAGCAAAACAG-----	629
Db	134	alileluginlagnluleuabserseglupheTySerargProgluleuProalProar	154
QY	630	----GATGTAGGATGACACACSTGGAAAGATGGACAGACSTTCAAGATGTGAG	685
Db	154	ghrsaprpserleuaseralleglnlntprapargthrtleuylgrspla--	173
QY	686	AAGCCAGGCTTTAAGGATTCAAATACAAAAATAAACAGATCTTTGGCACCTCC	745
Db	174	-----valylsglypheu-----leualaenpr	182
QY	746	TCGGCTCTCATCTACTCTCCCAAGTACCAACCTTACCCSAGCACACAGAGA	805
Db	182	oalalavalalvalthr-----Proglulnal	191
QY	806	GAGCAGCTACATCTTCGTCCAAAG-----	830
Db	191	aserclyalatyrglalearglualvalthrglyprovalglyleuvalleuglbr	211
QY	831	---CCACCTTTCAGAAAGTCCAGAGGGGGCCAGCAGAGAGATGCAAAAGACTTCAGT	889
Db	211	oprothleuproleunlsglnalaleuylthleuargrheleuleuglnllyseucluar	231
QY	890	GSTCTTGGACAGACAGAAATCTCACACACAGCAAAAGCGAATCTTTGGCCCATC	949
Db	231	gval-----	232
QY	950	ATCAACCAAAACACAGAAAGATGSCA---CGTCGATTCGC-----AGCTC	994
Db	233	-----alargalargalproserproalathrtalalvalnlylaleualaee	248
QY	995	TTCTCATGTCCA-----GAAAGCACAGTAT	1021
Db	248	ralapheglyproleuuleuatharglalarproproproglglylgnllyasglysercl	268
QY	1022	ACAAGCCAGACCATACAGTAGACATGACAGCACTGTCTGCTCAGAGATGCCAACTGC	1081
Db	268	uProalraprospheProvalleuleuLeuglualgrleuvalglnglnlyvalaspl	288
QY	1082	AGCCAGCAGCACCCCTCGAATGCTGCCCTTGAAGAACACAACTCGAAGAACTGAGCC	1141
Db	288	wglasprthalarproalaleupro--prolysproserlyvalylsyrProalar	307
QY	1142	CACA-----AAGCTGATGAAAGATGTCTGGCAGATGAATG	1186
Db	307	othralaleualaasnlylserthrtProserleuasnplala-----Gluthr	324
QY	1181	GTACATTGGAGAAATACGTGCGCAGGAGAGTGAAGATGTGTAATGAAGAGAACAGA	1241
Db	324	ptyltyrprolyasprlleserlrgeluglvalasnnglualgrleu--ArgasprthProas	343
QY	1241	TGGTACTTTTTTGGTCGAGACTGCTTCAAAATCAAGGAGACAACTATGTTTGGT	1300
Db	343	rglylthrheleuvalargasprlaserleuylleglnlgnllyu--Tutthleuth	362
QY	1301	GGCTTTTATGGCAACAAGTCTACATATGGAATAATCCCTTTCTCGAGACGATCAACA	1366
Db	362	rleuatyrgslglyasn-----asnlyseuileuylsvalrphenlsatdargsluNI	380

OY 1361 GTTGGCCCTGGGACAGACTAGAGAAATGAGATGTTGATCTGTGAGACATCAT 1420
 Db 380 STYGLYPHESERGLUPROLEU-----ThrPheCysSerValGluLeuI 396
 OY 1421 TGAACTACTACATAT 1436
 Db 396 ESERHISTYARGHIS 401

RESULT 4
 GRAP_HUMAN STANDARD; PRT; 217 AA.
 ID 013588:
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GRB2-related adaptor protein.
 GN GRAP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=96218119; PubMed=8647802;
 RA Feng G.-S., Ouyang Y.-B., Hu D.-P., Shi Z.-Q., Gentz R., Ni J.;
 RT "Grp is a novel SH3-SH2-SH3 adaptor protein that couples tyrosine
 RT kinases to the Ras pathway."
 RL J. Biol. Chem. 271:12129-12132(1996).
 CC -1- FUNCTION: COUPLE SIGNALS FROM RECEPTOR AND CYTOPLASMIC TYROSINE
 CC KINASES TO THE RAS SIGNALING PATHWAY.
 CC -1- SUBUNIT: ASSOCIATES THROUGH ITS SH2 DOMAIN WITH LIGAND-ACTIVATED
 CC RECEPTORS FOR STEM CELL FACTOR (KIT) AND ERYTHROPOIETIN (EPO).
 CC ALSO FORMS A STABLE COMPLEX WITH THE BCR-ABL ONCOPROTEIN. GRAP IS
 CC ASSOCIATED WITH THE RAS GUANINE NUCLEOTIDE EXCHANGE FACTOR SOS1,
 CC PRIMARILY THROUGH ITS N-TERMINAL SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE GRB2 / SEM-5 / DRK FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U52518; AAC50541.1; -
 DR HSSP; Q60631; IGRO.
 DR Genew: HGNC:4562; GRAP.
 DR MIM; 604330; -
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH3; 2.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR PRODOM; PD000066; SH3; 2.
 DR PRODOM; PD000093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 2.
 DR PROSITE; PS50001; SH2; 1.
 DR PROSITE; PS50002; SH3; 2.
 KW SH2 domain; SH3 domain; Repeat.
 FT DOMAIN 1 58 SH3 1.
 FT DOMAIN 60 152 SH2.
 FT DOMAIN 215 215 SH3 2.
 SQ SEQUENCE 217 AA; 25336 MM; 09PEC2F3BACDFAF8 CRC64;.

Alignment Scores: 0.0101 Length: 217
 Pred. No.: 146.50 Matches: 49
 Score: 44.83% Conservative: 29
 Percent Similarity: 44.83%

Best Local Similarity: 28.16% Mismatches: 48
 Query Match: 4.79% Indels: 48
 DB: 1 Gaps: 8

US-09-856-061-1 (1-1721) x GRAP_HUMAN (1-217)

OY 1131 AAACCTGACCCCAAGAGCTGATGAGAGATGTCTGCGAGATGATGATCATTTGA 1190
 Db 56 LysProHisPro-----TprTyrSerGly 63
 OY 1191 GAATACAGTCGCCAGGAGATGATGTGTTATGAAGAACAAAGATGCTACTTT 1250
 Db 64 ArgLysSerArgGlnLeuAlaGluGluLeuMetLysArgSerHisLeuGlyAlaPhe 83
 OY 1251 TTGTCGCGAGACCTCTTACAAATCCAGAGCAGAACCATATGTTTGGTGTTTAT 1310
 Db 84 LeuLeuArgGlu-----SerGluSerSerProGlyGluPheSerValSerValAsnTyr 101
 OY 1311 GGGAAAGGCTCTACATGTGCAAAATCCGTTTCTCGAGCAATCAACAGTTTCCCTG 1370
 Db 102 GlyAspGlnValGlnHisPheLysValLeuArgGluAlaSerGlyLysTyrPheLeu--- 120
 OY 1371 GGCACAGGACTACGAGAAATGAGATGTTGATCTGTGGAAGACATCATTTGAACACTAC 1430
 Db 121 -----TprGluGluLysPheAsnSerLeuAsnGluLeuValAspPheTyr 135
 OY 1431 ACATATTTTCCCATCTTCTGCTAATAGTGGAAACAAAGCTGCAGCCAGGAACAGTGC 1490
 Db 136 ArgThrThrThrIle-----AlaLysLysArgGlnIle 146
 OY 1491 TACCTC-----ACCAGCCACCTGCTC---TCCGACAGGCTCTTCTCAGTCACTACTCC 1541
 Db 147 PheLeuArgAspGluGluProLeu-LeuLysSerProGlyLysAla---CysPheAlaGlnAl 165
 OY 1542 AGCCAGGACCTTCATGACTAAGAAAGCCAGCCAGCATATC----- 1591
 Db 165 aglnPheAspPheSerAlaGlnAspProSerGlnLeuSerPheArgArgGlyAspIleI 185
 OY 1582 -----CCGATCAGTGGCTGCGG 1599
 Db 185 egluValLeuGluArgProAspProHisTprTprArgGly 198

RESULT 5
 P85B_BOVIN STANDARD; PRT; 724 AA.
 ID P85B_BOVIN
 AC P23726:
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 DE p85-beta subunit) (Ptdins-3-kinase p85-beta).
 GN PI3K2.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP MEDLINE=91191567; PubMed=1707345;
 RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Iarrea F., Panayotou G.,
 RA Thompson A., Dhand R., Hsuan J., Totley N., Smith A.D., Morgan S.J.,
 RA Courtneidge S.A., Parker P.J., Waterfield M.D.;
 RT "Characterization of two 85 kd proteins that associate with receptor
 RT tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase."
 RL Cell 65:91-104(1991).
 RN [2]
 RP CIRCULAR DICHOISM AND FLUORESCENCE SPECTROSCOPY.
 RX MEDLINE=93049176; PubMed=1330535;
 RA Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski B., Dhand R.,
 RA Fry M.J., Blundell T.L., Mollner A., Waterfield M.D.;
 RT "Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
 RT domain with a PDGF receptor phosphorylation site: structural features

Db	113	pleuProGluInPheSerProProAspValAlaProProIleuValLysLeuValGl	133
QY	608	CTCCACTGAGACACAACAGGAT-----GTGAG-----	638
Db	133	uAlaIleGluGthrGthrGlyLeuAspSerTyrArgProGluProProAlaIleAlaArgThrAs	153
QY	639	-----ATGACACAGCTGGAAAGAGGAGAACACCTCACTTCAGAGATGTCAAGAACCA	691
Db	153	prrPserLeuSerAspValGluGlnTrrPaspAlaIleAlaLeuSerAsp-----	169
QY	692	ACGCTTTAAAGGATTCAAATACACAAAATTAACMAAGACTCTTTGGCCACTTCGGCC	751
Db	170	-GlyAlaLysGlyPheLeu-----LeuAlaLeuProAlaPr	181
QY	752	TCCTATTCCT-----CTCCCAACAGTACCACCC	781
Db	181	oleuValAlThrProGluAlaAlaAlaGluAlaHisArgAlaLeuArgAlaAlaGlyPr	201
QY	782	CTTATCCCCACAGCA-----CCACCAGAGAGACAGCTACTTGGCTCCAA-----	828
Db	201	oValGlyProAlaLeuGluProPro-ThrLeuProLeuHisAlaLeuThrLeuAlaGr	221
QY	829	-----AGCCACACTT-----TCAGAAGTCCAGAGGGGCCAGCCAGAG	868
Db	221	heLeuLeuGln-HisLeuGlyArgValAlaGlyAlaArgAlaProAlaProGlyProAlaVal	240
QY	869	GAGTCAAAAGACTTCAGTAGGCTTCGGAGCAGAAAGAAATCTCAACACACACAA	928
Db	241	ArgAlaLeuGlyAlaThrPheGlyProLeuLeuArgAlaProProProPro-SerPr	260
QY	929	GCCAGAAATCTTCTGGCCATCATCAACCAACACACAGAGAGATCCACTGCCATTGC	988
Db	260	oProProGlyAlaLaProProspGlyThrGluProThrProAspPheProAlaLeuLeuVa	280
QY	989	CAGCTCTTCATACATGCCAGGAAAGCAACATATACAAAGCCAGACACTACAGTAGCAT	1048
Db	280	IGluLysLeuLeu-----GI	285
QY	1049	GCAGCACTGCTCCTCAGAGATGCCAAGCTGCAGCCACACGCCCTCGAATGCTGCC	1108
Db	285	ngLuhIserLeuGluGluGlnVal-----AlaProProAlaLeuPr	299
QY	1109	CTATGAAAACACAACACTCGAGAAACCTGACCCACAAAGCT-----	1151
Db	299	o-----ProLysProPolyThrLysProAlaProThrGlyLeuAl	313
QY	1152	-----GATGAGAAGATGTCGGCAGATGAATGTGATTCATTGGGAATACAGTCG	1201
Db	313	aAsnGlyGlySerProProSerLeuGlnsPalaGlnTrrPrrGlyAspIleSerAr	333
QY	1202	CCAGGCACTGCAAGATGTGTTAATGAAGAAGAACAAAGATGACTTTTGGTCCGAGA	1261
Db	333	gGluGluValAsnGluLysLeu---ArgAspThrProAspGlyThrPheLeuValArgAs	352
QY	1262	CTGCTTACAAAAATCCAAAGCAGCAACATATGTTTGGGGGCTTTTGGGAACAAGT	1321
Db	352	pAlaSerSerLysIleGlnGlyLys---TyrThrLeuThrLeuAlaGlyGlyGlyAsn--	370
QY	1382	ATCAGGAAATAGAGATGTTTGATTCCTGGAGAACATCAACATGTTGGCCCTGGCCAGAGACT	1381
Db	389	u-----ThrPheCysSerValAlaLysPheIleThrHisTyrArgHis	403

DE Tyrosine-protein kinase transforming protein fgr (EC 2.7.1.112).
GN V-FGR OR SRC-2.
OS Feline sarcoma virus (strain Gardner-Rasheed).
OC Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_TaxID=11775;
RN [1]
RX MEDLINE=84097512; PubMed=6318314;
RA Naharro G., Robbins K.C., Reddy E.P.;
RT "Gene product of v-figr onc: hybrid protein containing a portion of
RL actin and a tyrosine-specific protein kinase."; Science 223:63-66(1984).
CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FGR
CC POLYPEPTIDE.
CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC
CC SUBFAMILY.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X00255; CAA25063.1; -
CC PIR: A00653; TYMVR.
CC HSSP: P00523; 2PTR.
DR InterPro: IPR004001; Actin.
DR InterPro: IPR004000; Actin_like.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF00022; actin; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00268; ACTIN; 1.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS00406; ACTINS_1; 1.
DR PROSITE: PS01132; ACTINS_ACT_LIKE; 1.
DR PROSITE: PS50001; SH2; 1.
KW Polypeptide; Tyrosine-protein kinase; Transferase; Phosphorylation;
KW ATP-binding; Oncogene; SH2 domain.
FT DOMAIN 23 157
FT DOMAIN 167 264
FT DOMAIN 286 539
FT NF_BIND 292 300
FT BINDING 314 314
FT ACT_SITE 405 405
FT ACT_SITE 435 435
FT MOD_RES 435 435
SQ SEQUENCE 545 AA; 61516 MW; 7F4742EB7A7C413E CRC64;

Alignment Scores:
Pred. No.: 0.0125 Length: 545
Score: 146.00 Matches: 69
Percent Similarity: 35.27% Conservative: 28
Best local Similarity: 25.09% Mismatches: 101
Query Match: 4.78% Indels: 78
DB: 1 Gaps: 11

QY 747 GGGCTGCTATGCTCTGCCAAGAGTACCAACCTTACCCCGACACAGAGAG 806
|||
DB 6 ATGPTAALAAValCysArgProArgProLeuProProLeuProThAlaMetLuGl 25
QY 807 AGCAGTGCATPCTTC-----
|||
DB 26 GLUValAlaAlaLeuValIleAspAsnGlySerGlyMetCysLysAlaGlyPheAlaGly 45
QY 822 -----GCTCCAAAGCCCACTTTCCAGACAGAGGGGGCCCA----- 861
|||
DB 46 ASPAspAlaProArgAlaValAlaPheProSerIleVal-GlyAProArgHisGlnGlyVal 65
QY 862 -----GGCAGAGAGTGCCTCAAAAGACTTCAGAGAGGCTCTTGAGCAGAGAAGAAT-- 912
|||
DB 65 IMetValGlyMetGlyGlnLysAspSerTyrValGlyAspGlnAlaGlnSerLysArgI 85
QY 913 -----CTCACACACGACAAAGCCAGAA 935
|||
DB 85 YIleLeuThrLeuLysTyrProIleGlnHisGlyIleValIleThrAsnTyrPaspMetGl 105
QY 936 TCTTCTGCCCCATCATCAACCAACCAACAGACAGAGTCCACCTGCCAGCTCT 995
|||
DB 105 uLysIleTyrPHisHisThrPheTyrAsnGlnLeuArgVal----- 118
QY 996 TCTACATGCGCAGGAAGACAGATATACAGCCAGACACCATACAGTAGATGACGAC 1055
|||
DB 119 -----AlaProGlnGlu-----HisPr 124
QY 1056 TGTCTGCTCAGAGATGCCAAGCTGACGACGACGACGCTCGAATG-----CT 1105
|||
DB 124 OValleuLeuThrGluAlaProLeuAsnProLysAlaAsnArgGlnLysMetThrGlnI 144
QY 1106 GCCCTATGAAACACAAACCTCGGAGAACTGACCCCAACAAAGCCGTGAGAGAGATGT 1165
|||
DB 144 eMetPheGluThrPheAsnIle-----ProSerAsnTyrValAlaProValAspSerI 162
QY 1166 CTGCGCAATGATGTGTACATTTGAGAAATACAGTCCGACGAGGTGAA---GANTGT 1222
|||
DB 162 eGlnAlaGlnGluGlnTyrPheGlyLysIleGlyArgLysAspAlaGlnGlnLeu 182
QY 1223 AATGAAGAGACACAGAGAGTCTTTTGGTCCGAGACTGCTACAAATCAAGGC 1282
|||
DB 182 uSerProGlnLysAlaArgGlyAlaPheLeuValArgGlnSerGluThrLysGlnY 202
QY 1283 AGAACATATGTTTGGTGTG-----TTTATGAGCAACAGCTACAA 1327
|||
DB 202 A-----TyrSerLeuSerIleArgSerPirPaspGlnAlaArgGlyAspHisValLysH 220
QY 1328 TGTGAATCCGTTTCTCGAGAGCAATCAACAGTTTGGCCCTGGGACACAGACTACGAG 1387
|||
DB 220 sTyrLysIleArgLysLeuAspHrGly-----GlyTyrTyrIleThrTh 235
QY 1388 AATGAGATGTTGATCTGTGAGACATCAATGAACACTAC 1430
|||
DB 235 rArgAlaGlnPheAsnSerValGlnGlnLeuValGlnHisTyr 249

RESULT 7
P85B_HUMAN STANDARD; PRT; 728 AA.
ID P85B_HUMAN
AC 000459;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
DE p85-beta subunit) (Ptdins 3-kinase p85-beta).
GN PI3K2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RX SEQUENCE FROM N.A. PubMed=9582025;
MEDLINE=98241181;

DR	EMBL: X80907; CAA56868.1; -
DR	HSSP: P23727; 2PNB.
DR	Genew: HGNC:8980; PIK3R2.
DR	MIM: 603157; -
DR	InterPro: IPR001720; PI3kinase_P85.
DR	InterPro: IPR000198; RHOGAP.
DR	InterPro: IPR000980; SH2.
DR	InterPro: IPR001452; SH3.
DR	Pfam: PR00017; SH2; 2.
DR	Pfam: PR00018; SH3; 1.
DR	Pfam: PF00620; RHOGAP; 1.
DR	PRINTS: PR00678; PI3KINASEP85.
DR	PRINTS: PR00401; SH2DOMAIN.
DR	ProDom: PD000066; SH3; 1.
DR	ProDom: PD000093; SH2; 2.
DR	SMART: SM00324; RHOGAP; 1.
DR	SMART: SM00252; SH2; 2.
DR	SMART: SM00326; SH3; 1.
DR	PROSITE: PS50001; SH2; 2.
DR	PROSITE: PS50002; SH3; 1.
KW	SH3 domain; SH2 domain; Repeat.
FT	DOMAIN 4 80 SH3.
FT	DOMAIN 125 258 RHG-GAP.
FT	DOMAIN 330 425 SH2 1.
FT	DOMAIN 622 716 SH2 2.
SEQUENCE	728 AA; 81624 MW; 80C2AE244977346B CRC64;

QY	390	CCACITCCGGATCACAAGAGAACTTGGCTGGGGGCTCCAGTGGAGAAATAATGCACAGT	449
Db	95	ProLeuP ^{ro} Ala ^{tr} Pro ^{tr} Pro ^{tr} Gly ^{tr} Ala ^{tr} Pro ^{tr} Glu ^{tr} Leu ^{tr} Thr ^{tr} Leu ^{tr} Pro ^{tr} S ^{tr} Leu	114
QY	450	AACAAACGCTACGAGATCCTGAGTTCAG-----CTGCTAAGGATCGGCATCA	500
		:::	
Db	115	ProGlu ^{tr} Gln ^{tr} Phe ^{tr} Ser ^{tr} Pro ^{tr} Ala ^{tr} Val ^{tr} Ala ^{tr} Pro ^{tr} Leu ^{tr} Val ^{tr} Leu ^{tr} Val ^{tr} Glu ^{tr} Ala	134
		:::	
QY	501	ATGAAA-----ATTATACAGCCAGCATATCCAGAA	553
Db	135	IleGlu ^{tr} Ala ^{tr} Gly ^{tr} Thr ^{tr} Gly ^{tr} Leu ^{tr} Ser ^{tr} Gln ^{tr} Ser ^{tr} His ^{tr} Tyr ^{tr} Ala ^{tr} Pro ^{tr} Glu ^{tr} Ala ^{tr} Pro ^{tr} Arg	154
QY	534	TGCGAATACGAA-----GATACACGCTATTTCACAGATATGATG	572
		:::	

Db	155	ThirapserleuserAspValaSpolnTripsaphrhalalaLeuAlaAspGlyLe	174
QY	573	GAGGCTCCCTTCG---TTACST-----CCAGGCTTCGTCTCCACT	614
Db	175	LyssSerPheLeuAlaLeuPheProValaThrProGluAlaSerAlaGluAla	194
QY	615	GAGACACAAACCAGGATGTGAGATGACACAGCTGSAAGACTGGACAAAGCTTACTTC	674
Db	195	ArgatgAlaLeuAlaGgGluAlaAlaGlyProValaGlyProAlaLeuGluProProthLeu	214
QY	675	AAGGATGTCAGAACCCCAACGCTTTAAAGATTCAAAATACAAAAATAAACAGACTCCT	734
Db	215	ProLeuHisArgAlaLeuThrLeuArg--PheLeuLeuHisLeuGlyArgVala	233
QY	735	TTGGACCTCCCTGG-----CCGGCATC---ACCTCCCAAGAAAGTACCAACCTTA	785
Db	234	ArgatgAlaProAlaLeuGlyProAlaValaArgAlaLeuGlyAlaThrPheGlyProLeu	253
QY	786	-----CCCCAGACACACACAGAGAGAGAGAGATGTCATCTCCCTCCAAAGGCCACC	836
Db	254	LeuLeuArgAlaProProProProSerSerProProGlyGlyLysAlaProAspGlySer	273
QY	837	TTTCCAGAAAGTCCAGAGGGGGCCACGACAGAGAGTGCACAAAGACTTCACTAGGCTCT	896
Db	274	GluPro-----SerProAspPheProAlaLeuLeu	283
QY	897	GGAGCAGAGAAAGAAATCTCACACACAGACAAACCCAAAGCTTCCTCCGATCATCAAC	956
Db	284	ValGlyLysLeuLeuGlnGlnHisLeuGlnGlnGlnValaLysProAlaLeuPro	303
QY	957	CAAAACACACAGAAAGACTCCACTGCCATTGCCAGCTTCTCCATCATGCCAGAAAGAC	1016
Db	304	ProLysProProLysAlaLysProAla-----	312
QY	1017	AGATTACAAACCCAGACACATACAGAGACATGCACAGCTGCTGCTGCTCAGAGATGCCA	1076
Db	312	-----	312
QY	1077	GCTGCAGCCAGACACACACCTCTGAAATGCTGCCCTATGAACAAACATCGAGAAACT	1138
Db	313	-----ProThrValLeu-----AlaAsnGlyLysSerPro	322
QY	1137	GACCCACAAAGCCTGATGGAAGAGATGCTGCGAAGATGATGGTACATTGGCAAAATAC	1196
Db	323	-----ProSerLeuGlnAspAla-----GluThrPyrTrpLysPhe	335
QY	1197	AGTCCAGGACGAGTGAAGATGTGTTAATGAAGAAGACACAGATGTAATTTTGGTC	1256
Db	336	SerArgGlnGlnValAlaAsnGlnLysLeu--ArgAspThrProAspGlyThrPheLeuVal	354
QY	1257	CGAGACTGCTCTCAAAATCCAAAGGACGACAAACATATGTTTGGTGGCTGTTTATGGGAC	1316
Db	355	ArgAspAlaSerSerLysIleGlnGlyGlu--TyrThrLeuThrLeuArgLysGlyGly	373
QY	1317	AAGGTCACAAATGGAATTCGTTCTCTCGAAGACAAATCAACATGTTTGGCTGGGACACA	1376
Db	374	Asn-----AsnLysLeuIleLysValPheHisArgAspGlyHisLysGlyPheSerGly	391
QY	1377	GGACTAGAGGAATAGATGTTTGAATTTGTGTGGAAAGACATGTAACACATACATAT	1436
Db	392	ProLeu-----ThrPheCysSerValValAspLeuIleAsnHisLysTyrArgHis	407
RESULT 8			
SCK_HUMAN	STANDARD:	PRT:	540 AA.
ID	SCK_HUMAN		
AC	P98077; O9NP15; O60230;		
DT	01-FEB-1996 (Rel. 33, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Protein SCK (Fragment).		
GN	SCK.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		

CC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX	NCBI_TaxID=9606;
RA	SEQUENCE FROM N.A.
RP	MEDLINE=98175964; Pubmed=9507002;
RX	Nakamura T., Murakawa S., Sanokawa R., Mori N.;
RT	"Shc and Sck, two neuronally expressed Shc adapter homologs. Their
RT	differential regional expression in the brain and roles in
RT	neurotrophin and Src signaling."
RL	J. Biol. Chem. 273:6960-6967(1998).
RN	[2]
RP	SEQUENCE OF 72-499 FROM N.A.
RX	MEDLINE=95090462; Pubmed=7527937;
RA	Kavanaugh W.M., Williams L.T.;
RT	"An alternative to SH2 domains for binding tyrosine-phosphorylated
RT	proteins."
RL	Science 266:1862-1865(1994).
RN	[3]
RP	SEQUENCE OF 72-328 FROM N.A.
RA	Immerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA	Burkhardt-Schultz K., Gordon L., Dias J., Brower A., Ramirez M.,
RA	Stilwagen S., Phan H., Velasco N., Do L., Regula W., Terry A.,
RA	Garnes J., Dangnan L., Erler A., Christensen M., Georgescu A.,
RA	Avila J., Liu S., Attix C., Andreise T., Trankhelm M.,
RA	Amico-Keller G., Coefield J., Duarte S., Lucas S., Sanders R.,
RA	Thomas P., Quan G., Kronmiller B., Arellano A., Sanders C., Ow D.,
RA	Nolan M., Trong S., Kobayashi A., Olsen A.S., Carraro A.V.;
RT	"Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a
RT	serine protease gene cluster."
RL	Submitted (DEC-1998) to the EMBL/Genbank/DBJ databases.
RN	[4]
RP	SEQUENCE OF 149-540 FROM N.A.
RA	Carim L., Estivill X., Sunoy M., Escarceller M.;
RL	Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.
CC	-1- TISSUE SPECIFICITY: PREDOMINANTLY IN LIVER, ALSO PRESENT IN BRAIN.
CC	-1- SIMILARITY: CONTAINS 1 PID DOMAIN.
CC	-1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC	-1- SIMILARITY: TO SHC TRANSFORMING PROTEINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC	or send an email to license@isb-sib.ch).
CC	-----
DR	EMBL; AB001451; BAA25798.1; -
DR	EMBL; AC006124; AAC37988.1; -
DR	EMBL; AL360234; CAB96175.1; -
DR	HSSP; P29353; IMTL.
DR	MIM; 605217; -
DR	InterPro; IPR000050; PID_domain.
DR	InterPro; IPR000980; SH2.
DR	Pfam; PF006407; SH2; 1.
DR	Pfam; PF00640; PID; 1.
DR	PRINTS; PR00401; SH2DOMAIN.
DR	PRINTS; PR00629; SHCPIDOMAIN.
DR	ProDom; PD000093; SH2; 1.
DR	SMART; SM00462; PTB; 1.
DR	SMART; SM00252; SH2; 1.
DR	PROSITE; PS01179; PID; 1.
DR	PROSITE; PS0001; SH2; 1.
DR	SH2 domain.
KW	
FT	NON_TER 1
FT	DOMAIN 105 287 PID.
FT	DOMAIN 445 536 SH2.
FT	CONFLICT 72 72 A -> P (IN REF. 2).
SO	SEQUENCE 540 AA; 57623 MW; 0DB3700910B1A5DB CRC64;

Alignment Scores: 0.0297 Length: 540

Pred. No.: 140.50 Matches: 101

Percent Similarity:	34.59%	Conservative:	55
Best Local Similarity:	22.39%	Mismatches:	196
Query Match:	4.60%	Indels:	99
DB:	1	Gaps:	19
US-09-856-061-1 (1-1721) x SCK_HUMAN (1-540)			
QY 243	GTGCCAAGCAGCATGACACACGCCAGGCATTAAGACACAGAAAGAA-----		290
Db	148 ValProGlyValAlaArgGlySerThrPylsLysAlaProAsnLysAlaLeuAlaSerVal		167
QY 291	---GGATTCCGGTATCTGACATCTCCGAACGTCCTCTGCGTGAATAATAGTCATGGCCA		347
Db	168 LeuLysSerAsnLeuAlaArgPheAlaGlyMetSerLleSerLleHisLleSerThrAsp		187
QY 348	AGCCTCAGCAGTGTCCCAAGAGGCGGTCCAGCGGTTCTTGAAACCACTCCGGATCACAGA		407
Db	188 GlyLeuSerLleLeuSerValProAlaThrArgGlnValIle-----AlaAsnHisLMet		205
QY 408	AGGAACCTTGCTGGGGTCCAGCTGGAGAAATAACACAGTACACACAGTAC-----		461
Db	206 ProSerLleSerPheAlaSerGlyAsp-----ThrAspMetThrAspTyrValAla		223
QY 462	-----GAAGTCTCTGAGTTCCAG-----CTGCTCAAGAGCATGGCCATCA		500
Db	224 TyrValAlaLysAspProLleAsnGlnAlaArgAlaCysHisLleLeuGlnLysCysGlnGly		243
QY 501	ATGAAATTTTACCAGCCAGCACACCTTCCAGGAATCG-----GAATTCAGCATACAGC		554
Db	244 Leu-----AlaGlnSerLleLleSerThrValGlnGlnAlaPheGlnLeuAlaArg		259
QY 555	TATTTCCAGATATGATGAGAGCTCC--CTTCTGTATCTCCCAAGGCTTGTCTCC		611
Db	260 PheLysGlnTyrLleHisSerProProLysValAlaLeuProGlnLysAlaGly		279
QY 612	ACTGAGACACAAACCCAGGAGTGTGAGAGTACACACCTGGAA-----		653
Db	280 ProGlnGlnSerAlaTyrGlyAspArgGlnAspSerLeuGlnHisAsnTyrTyrAsnSer		299
QY 654	-----GAAGTGCACAAAGCCTTCCAGGATGTCCAGAAAGCCAAAGCCTTTAA		701
Db	300 IleProGlnLysGlnProProLeuGlnGlyLeuValAspSerThrGluAlaLeuThrGln		319
QY 702	GGATTCAATACACAAATAAACACAGACTCTTTGCCACCTCCGCGCTATCTACT		761
Db	320 ProCysAlaLeuThrAlaLeuAspGlnGlyProSerProSerLeuAlaGspAlaCysSer		339
QY 762	CTCCCCAAAGATACCAACCTTACCCGCCACACACAGAGAGACAGTGCATCTCC		821
Db	340 LeuProThrAspValAlaGlySerThrGlyThrAlaProGlnLysAspLys-----		355
QY 822	GCTCCAAAGCCCACTTCCAGAACTCCAG-----AGGGGGCCCAAGCAGAGAGT		872
Db	356-----TyrValGlnAlaAspAlaArgGlyPro-----		364
QY 873	GCMAAAGACTTCAGTAGGGCTCTTGAGAGCAAGAAAGATCTCCACCACACAAAGGCCA		932
Db	365---ProAspHisGlnGlnHisLeuTyrValAsnThrGlnGlyLeuAspAlaProGlnPro		383
QY 933	GAATCTTTCTGGCCCATCTCAACACCAAAACACACAGAAAGTGCACCTGCCATTGCCAGC		992
Db	384 GluAspSer-----		386
QY 993	TCTTCTCCATATGCCAGAAAGCAGATATACAAAGCCAGAGACCATACAGGTAGCATGCCAG		1052
Db	387-----ProLysLysAspLeuPheAspMetArgProPheGlnLysPheAlaLeuLys		402
QY 1053	CACTGTCTCTCAGAGATGCCAAGTGCACAGCCACAGCCACCTTCGAATCTGCCCTAT		1112
Db	403-----LeuHisGlnLysCysValAlaAlaGlyValAlaThrAlaLeuProLeuProLeu		419
QY 1113	GAAAACACAACTCGGAGAAACCTGCAC-----CCACAAAGCTCTATGAGAAAGAT		1169

DB 420 GUAAGSINTPRPProSerProBiotharArgAlaProValAlaProThylGluIn 439
 QY 1164 GCTGCGAGATGATGCTACTTGGAGAAATACAGTCCGACGAGAGATGCTTA 1223
 DB 440 LeuAAGGInGluProThylArgHisGlyArgMetSerArgAlaAlaGluArgMetLeu 459
 QY 1224 ATGAAGAACAACAGATGCTTTGGTGGAGAGAGTCTACAAATGTAATCCAGCA 1283
 DB 460 -----ArgAlaAspGlyAspPheLeuValArgAspSerValTh-----AsnPro 474
 QY 1284 GAACCAATGTTTGGTGGTCTTTATGGAACAAGCTCTACAAATGTAATCCGTTTC 1343
 DB 475 GlyIntYrValLeuThGlyMetHisAlaGlyInProLysHisLeuLeuValAsp 494
 QY 1344 CTCGAGACCAATCAACACTTTGCCCTGGCACAGACACAGAGAAATGATGTTGAT 1403
 DB 495 ProGlu-----GlyValAlaArgThrLysAspValLeuPheGlu 507
 QY 1404 TCTGTGAAGACATCATTTGAACACTACACATATTTCCCATTCGCTAATGATGGAAA 1463
 DB 508 SerIleSerHisLeuIleAspHisHis-----LeuGlnAsnGlyIn 521
 QY 1464 GACAAGCTGCAGCAGCAAGAACAGTGCACCTC 1496
 DB 522 ProIleValAlaAlaGluSerGluLeuHisLeu 532

RESULT 9
 TENS_CHICK STANDARD; PRT: 1744 AA.
 AC 004205: 091007: 092011:
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 01-NOV-1997 (Rel. 35, Last annotation update)
 DE Tensin.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OC NCBI_TaxID=9031;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=94350987; PubMed=8071358;
 RA Lo S.H., An Q., Bao S., Wong W.K., Liu Y., Janney P.A., Hartwig J.H.,
 Chen L.B.;
 RT "Molecular cloning of chick cardiac muscle tensin. Full-length cDNA
 RT sequence, expression, and characterization."
 RL J. Biol. Chem. 269:22310-22319(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Heart;
 RX MEDLINE=95204530; PubMed=7896874;
 RA Chuang J.Z., Lin D.C., Lin S.;
 RT "Molecular cloning, expression, and mapping of the high affinity
 RT actin-capping domain of chicken cardiac tensin."
 RL J. Cell Biol. 128:11095-1109(1995).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chen L.B.;
 RL Submitted (XX-1991) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE OF 1469-1744 FROM N.A.
 RC TISSUE=Embryonic chondrocytes, and Embryonic heart;
 RX MEDLINE=94039118; PubMed=823621;
 RA van de Werken R., Gennari M., Tavella S., Bet P., Molina F.,
 Lin S., Canceda R., Castagnola P.;
 RT "Modulation of tensin and vimentin expression in chick embryo
 RT developing cartilage and cultured differentiating chondrocytes."
 RL Eur. J. Biochem. 217:781-790(1993).
 RN [5]
 RP SH2 DOMAIN.
 RX MEDLINE=91220073; PubMed=1708917;
 RA Davis S., Lu M.L., Lo S.H., Lin S., Butler J.A., Druker B.J.,

RA Roberts T.M., An Q., Chen L.B.;
 RT "Presence of an SH2 domain in the actin-binding protein tensin."
 RL Science 252:712-715(1991).
 CC -1- FUNCTION: MAY BE INVOLVED IN CELL MIGRATION, CARTILAGE DEVELOPMENT
 CC AND IN LINKING SIGNAL TRANSDUCTION PATHWAYS TO THE CYTOSKELETON..
 CC BINDS TO ACTINS AND PHOSPHORYLATED PROTEINS IN SRC-TRANSFORMED
 CC CELLS. MAY BIND ACTIN WITH CAPING AND BUNDLING PROPERTIES.
 CC -1- SUBCELLULAR LOCATION: LOCALIZES TO ADHERENS JUNCTIONS.
 CC -1- TISSUE SPECIFICITY: HEART, GIZZARD, LUNG AND SKELETAL MUSCLE.
 CC -1- PTM: TYROSINE-PHOSPHORYLATED.
 CC -1- SIMILARITY: CONTAINS 1 TENSIN DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----
 DR EMBL: M96632; AAA59053.1; -;
 DR EMBL: L06662; AAA73949.1; ALT_INIT.
 DR EMBL: Z18529; CAA79215.1; ALT_INIT.
 DR EMBL: M74165; AAA49087.1; -;
 DR EMBL: X66286; CAA46992.1; -;
 DR HSP: P16277; 1BLK.
 DR InterPro: IPR000340; DS_phosphatase.
 DR InterPro: IPR000050; PID_domain.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00462; PTB; 1.
 DR SMART: SM00012; PTBc_DSPc; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 KW Actin-binding; Cytoskeleton; SH2 domain; Phosphorylation.
 FT DOMAIN 66 342
 FT FT 1472 1581
 FT CONFLICT 49 49
 FT CONFLICT 61 61
 FT CONFLICT 88 88
 FT CONFLICT 404 404
 FT CONFLICT 452 452
 FT CONFLICT 508 509
 FT CONFLICT 522 522
 FT CONFLICT 664 664
 FT CONFLICT 666 666
 FT CONFLICT 875 875
 FT CONFLICT 909 909
 FT CONFLICT 1102 1113
 FT CONFLICT 1240 1240
 FT CONFLICT 1480 1480
 FT CONFLICT 1711 1711
 SQ SEQUENCE 1744 AA; 187214 MW; 5C3CB6211935524 CRC64;
 Alignment Scores:
 Pred. No.: 0.0377
 Score: 140.00
 Percent Similarity: 34.11%
 Best Local Similarity: 21.58%
 Query Match: 4.58%
 DB: 1
 Gaps: 21
 US-09-856-061-1 (1-1721) x TENS_CHICK (1-1744)
 QY 527 CCAGGATGGAATGAGATACAGTATTTCCAGATATGAGAGTCCCTTCT 586
 DB 1299 ProGluLeuGly-----ArgHisThrValSerProHisAlaProProGlySerProSer 1316
 QY 587 GTTACTCCCAA-----GGTTTGTGTCTC 610
 DB 1317 LeuAlaArgHisGluMetAlaAlaValProProGlySerProMetGlycyltyrSer-Se 1336


```

Query Match: 4.56# Indels: 118
DB: 1 Gaps: 20
US-09-856-061-1 (1-1721) x PSC_DROME (1-1603)

QY 278 GACAAAGAAAGAAAGATTCGTGATTCGATTCGAAACGCTCTCTGCTGAAATAAG 337
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 613 AAPPProgluArgArg-----GluIleVallyProleuLysProgluLysGlu 628
QY 338 GTCATGGCCAAAGCTCAGAG-----TGCAGAAAGGCGGTTCGAGCGTTCTGA 388
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 629 SerArgSerLysLysLysAspLysAspLysArgLysProLysSerSerSerSerSer 648
QY 389 ACCACTTCGCGATCAGAGAAAGAACTGGCTGGCTGCTCC----- 427
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 649 SerSerSerGlyGluArgLysArgLysSerProSerProLeuThrValProLeuThr 668
QY 428 ---AGGTGAGAAATAATCAACAGTACACAGACTAC----- 461
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 669 IleArg-ThrGluArgIleMetSerProSerGlyValSerThrLeuSerProArgValTh 688
QY 462 -----GAAGATCCCTCGAGTTCAGCTGCTGAAGCATGCGCA----- 497
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 688 rSerGlyAlaPheSerGluAspProLysSerGluPheLeuSerPheAlaLeuLysPr 708
QY 498 -----TCAATGAATAATTTTACAGC 517
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 708 oLleLysVallyValLysUserProGluArgThrLeuAsnAsnArgAlaIleThrProPr 728
QY 518 CAGACCT---ATCCAGAAATGGAAATACGACATACAGC-----TATTTCAGGAT-- 566
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 728 oSerProSerValGlnGlnSerAlaSerProLysSerLysLysAsnLeuAspAspSe 748
QY 567 ---ATGATGAGAGGCTCCCTTGTTCCTCCAGGCTTCTGCTCCACTGAGAGCA 622
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 748 rIleLeuMetLysProProSerCysMetProProLysSerIleAlaSerLysArgLys 768
QY 623 AACCAAGGAT-----GTGAGAGTACACAGCTGAGAGAAAGTGACAGACCTAC 670
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 768 sSerLysGluProValLysAlaValSerLysLysGlnLysLeuSerProProLeuProTh 788
QY 621 C-----TTTCAG-----GATGTCAGAAAGCCAGCTTTAAAGATCAATAACAC 715
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 788 rValAspPheLysIleArgLysProValThrAsnGlyAsnSerSerGlyThrAlaSerPr 808
QY 716 AAAATAAACAAGACTCTTGTCCACCTCTCT---CGGCGCTGTCTCACTGCCCAAGAA 772
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 808 oLysIleGluLysProLeuMetProProProAlaLysProMetLeuAlaProArgLys 828
QY 773 GTACCAACCTTA-----CCCCAGACACACAGAGAGAGAGACTGCA-- 815
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 828 sLeuGlnProSerAlaGlnPheAlaProProProSerProIleHisHisAlaGlyVa 848
QY 816 ---TACTTGCCTCCA-----AAGCCCACTTTTC 841
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 848 lGlnMetSerAlaProLysAsnArgThrProIleAlaLysArgThrGlnProIleLeuPr 868
QY 842 AGAAGTCCAGAG---GGGCCAGAGAGAGAGAGTGCAGAAAGACTTCAGTGGGCTCT-- 896
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 868 oLysAlaSerArgProAsnProPheAlaAsnIleProAsnAspValAsnArgLeuLys 888
QY 897 -----GGAGCAGAGAGAAAGAACTCAGCA 919
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 888 sAspAlaGlyThrGluIleLysSerIleGlyGlySerValGlnLysAsnSerAsnSe 908
QY 920 CCAGACAAACCCAGAACTCTTGGCCATATCAACAAACAAACACACAGAGAGTCCACC 979
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 908 rAlaGlnLysProHisLeuThrGlyProLysGly-----GluThrLysMetGlyProp 926
QY 980 TGGCATGTCAGGCTCTTCATACATGCAGAAAGACAGTATACAAAGCCAGACCATAC 1039
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926 oAlaLeuProAlaThrThrProSerGlnGlyAsnLysAsnValGlyLysGlnAlaLys 946

```

```

QY 1040 AGTAGCATGACGACACTGCTCT----- 1061
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 946 nLeuProMetSerAlaProProAsnLysGlyAsnSerSerAsnAsnThrLeuAla 966
QY 1062 -----GCTCAGAAATGCCAAAGCTGACAGCCAGACAGCCCTCGAATGCTGCCCTA 1111
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 966 aLeuPheAsnSerAsnLysCysLysGlyLysGlnAlaProProGlyCysArgThrProMe 986
QY 1112 TGAAMAACAAACCTCGGAGAAAGCTGACCCAGAACCCCTGATGAGAGATGCTGGCA 1171
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 986 tTyThrProAsnSerProIleTySerProSerSerProLys----- 1000
QY 1172 GAATGATGCTACATTTGGAGATACACT 1199
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1001 -----TyThrValProSerTyThrAsn 1006

RESULT 11
ABP1_YEAST STANDARD: PRT: 592 AA.
AC P15891;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Actin binding protein.
GN ABP1 OR YCR088W OR YCR88W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90136906; PubMed=2405279;
RA Drubin D.G., Mulholland J., Zhu Z., Botstein D.;
RT "Homology of a yeast actin-binding protein to signal transduction
RN Nature 343:288-290(1990).
RL [2]
RP SEQUENCE FROM N.A.
RA Frontali L., Grisanti P.;
RL Submitted (MAR-1992) to the EMBL/GenBank/DBJ databases.
CC - FUNCTION: MAY BE INVOLVED IN THE SPATIAL ORGANIZATION OF CELL
CC SURFACE GROWTH. AN OVERPRODUCTION OF ABP1 CAUSES THE ASSEMBLY OF
CC THE CORTICAL ACTIN SKELETON AT INAPPROPRIATE SITES ON THE CELL
CC SURFACE, RESULTING IN DELocalized SURFACE GROWTH.
CC - SUBCELLULAR LOCATION: CORTICAL CYTOSKELETON.
CC - SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC - SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X51780; CAA36075.1; -
DR EMBL: X59720; CAA42253.1; -
DR PIR: S19503; LBLY.
DR HSSP: P29355; ISEM.
DR SGD: S0000684; ABP1.
DR InterPro: IPR002108; Actbind.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00241; cofilin_ADF; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR PRODOM: PD000066; SH3; 1.
DR SMART: SM00102; ADF; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
DR Cytoskeleton; Actin-binding; SH3 domain; Repeat.
KW TO YEAST COFILIN.
FT SIMILAR 1 142
FT DOMAIN 86 97 ACTIN-BINDING (POTENTIAL).

```

```

FT DOMAIN 532 592 SH3
FT REPEAT 200 575 3 X 10 AA APPROXIMATE REPEATS.
FT REPEAT 200 209 1
FT REPEAT 436 445 2
FT REPEAT 566 575 3
FT CONFLICT 58 58 L -> S (IN REF. 1).
FT CONFLICT 312 312 K -> I (IN REF. 1).
SO SEQUENCE 592 AA: 65576 MW: 39523510704D94AA CRC64:

Alignment Scores:
Pred. No.: 0.0482 Length: 592
Score: 137.50 Matches: 84
Percent Similarity: 36.88% Conservative: 58
Best Local Similarity: 21.82% Mismatches: 135
Query Match: 4.50% Indels: 108
DB: Gaps: 20

US-09-856-061-1 (1-1721) x ABPI_YEAST (1-592)
QY 221 GAAGATCCCTCTCTTTGGCAGGTGCGCAGGACCATGACCGAGGCAATAAAGGAC 280
DB 244 GIUASPPHrgrleuVal-GIuysPro-----ThraIaIaGIySerIySileas 260
QY 281 AACGAAGAAGATTCGTGATCTGAGATTCAGAAAGCTCTCTGCTGAAAAATAGTTC 340
DB 260 pProSerSerAspIlaIaIaSnleuys-----AsnGIuSerIySleuylsArgaspse 278
QY 341 ATGGCCAGCCTCAGCAGTCCCAAGGCGGTTCGAGCGGTTCGAGACCATCTCCGGA 400
DB 278 r-----GIuPhaSnSerPheluGIyThrThIySPro----- 289
QY 401 TCACAGAGAACTTGGCTGGGTCCAGGTGAGAAATAATGCAACAGTAACAGCATA 460
DB 290 -----ProSerMetThrGIuSerSerIeulysSnasp-- 300
QY 461 CGAAGATCCTGATCCAGCTGCTGGAAGCATG-----CCATCAATGAA 505
DB 301 -----AspAspIyValIleIySgIyPhaArgSnGIuIySerProIaGIuIe 317
QY 506 AATTTTACCGCAGACCTTCCAGAGATGGAATACGCAATACGCAATACGCAATTTCCGGA 565
DB 317 UTTPaIaGIuArgIySaIaIySgIaSnSerGIyAsnIaGIuThrIyS----- 333
QY 566 TATGATGAGGCTCCCTCTCTTACCTCCCAAGGCTTCTCTCCAGTACAGACAAAC 625
DB 334 -----AlaGIuIaIa-----ProIySProGIuValProGIuSnp----- 344
QY 626 CAGGATGTGAGATGACACAGCTGGAAGAAGTGCACAGCTTCAAGATGTTCAG 685
DB 345 -----GIuProGIuGIuGIuProAspValIySAspIeulY 356
QY 686 AAGCAAGCTTTAAGGATTCAAATACAAAAATAAACAAGCTCTTGG----- 737
DB 356 sSer---LysPhaGIuIyIeuaIaIaIaSerGIuIySgIuGIuIeulY 375
QY 738 -----CCACCTCTCTCG-----CCTGTATCATCTCTCCCAAGAAGTAC--CA 778
DB 375 sPhaIaIaProProIySgIySserGIuProThIleIleSerProIySProPheSerly 395
QY 779 ACCCTTACCCCGACACACAGAGAGAGAGAGATGCTACTCTCCCAAGCCACTT 838
DB 395 sPProGIuGIuProValIySaIaGIuIaGIuIa-----GIuProIySThrAspTY 412
QY 839 TCCAGATTCAGAGGGGCCCGACGAGAGAGATGCAAAAGACTTCAGTAGCTCTTGG 898
DB 412 rIySlySlyIleIy-----AsnProIeulProGI 421
QY 899 AGCAGAAGAAGATCTCAGCAGACAGAAACCAAGATCTTCT----- 941
DB 421 yMetHIsIleGIuIaIaIaSpaSnGIuGIuGIuProGIuIaSnspAspAspAsp 441
QY 942 -----TGCCCATCATCAAC----- 956

```

```

DB 441 paspGIuSgIuIaIaIaGIuIaProProIeulProSerArgAsnValIaIaSerGIyIaIaPr 461
QY 957 -----CAAAACACACAGAAAGATGCCACCTCGCATTCGACCTCTTACATGCGAGAA 1012
DB 461 oValGIuIySgIuGIuGIuGIuGIuGIuGIuIaIaIaProSer-----LeuProSerAr 479
QY 1013 GCACAGTATACAGACCGACGAGACCATPACAGTGCATGACGACGACTCTCTGTCAGAGATG 1072
DB 479 gAsnSerIleProIaIaProIySgIuGIuIaIaProGIuIaIaProGIuIaIaProGIuIuI 499
QY 1073 CCAAGCTGCAGCCAGCCAGACCCCTGGAATGCTGCCCTTGAACACAACTCGAGAA 1132
DB 499 eGIuGIuGIuIaIaGIuGIuIaIaIaProGIuIeulProSerArgSerIaIaIaProPr 519
QY 1133 ACCTGACCCCAAAAG-----CCTGATGAGAAGATGTCTGCGAAGATGAATGATCATTTG 1189
DB 519 oProProIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 539
QY 1190 AGAATPACAGTCCGACGAGTGAAGATGTATGAAGAAGAACAGAGATGATCTTT 1249
DB 539 aGIuTyraSPTyraSPTyraSPTyraSPTyraSPTyraSPTyraSPTyraSPTyraSPTy 553
QY 1250 TTGTGTCGAGAC 1262
DB 553 eValGIuAsnAsp 557

RESULT 12
P85B_MOUSE
ID P85B_MOUSE STANDARD: PRT; 722 AA.
AC 008908;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
DE p85beta subunit) (Ptdins-3-kinase p85-beta).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
RX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH SWISS;
RC MEDLINE=98241181; PubMed=9582025;
RA Janssen J.W.G., Schleithoff L., Bartram C.R., Schulz A.S.;
RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase
RT p85beta subunit and HMOX8, a putative deubiquitinating enzyme.";
RL Oncogene 16:1767-1772(1998).
CC -I- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC -I- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC -I- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -I- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -I- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -I- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DB EMBL: Y13569; CAJ73903.1;
DB HSP: P23727; ZPNB.
DB MGD: MGI:1098772; PIK3R2.
DB InterPro: IPR001720; PI3kinase_P85.
DB InterPro: IPR000198; RHOgap.
DB InterPro: IPR000980; SH2.
DB InterPro: IPR001452; SH3.

```

DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00678; PIKINASEP5.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000066; SH3; 1.
 DR PRODOM: PD000093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 KW SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 4 80 SH3.
 FT DOMAIN 125 257 RHO-GAP.
 FT DOMAIN 324 419 SH2 1.
 FT DOMAIN 616 710 SH2 2.
 SQ SEQUENCE 722 AA; 81251 MW; AB4D49AF30CEC567 CRC64;

Alignment Scores:
 Pred. No.: 0.0495 Length: 722
 Score: 137.50 Matches: 102
 Percent Similarity: 36.498 Conservative: 52
 Best Local Similarity: 24.178 Mismatches: 151
 Query Match: 4.50% Indels: 117
 DB: 1 Gaps: 25

US-09-856-061-1 (1-1721) x P85B_MOUSE (1-722)

QY 365 AGGCGGGTGTGCGAGCGGTTCTGGAACCACTCCGATCCAGACAG-----GACTT 415
 DB 35 ATGVALALALALeuGlnALALeuGlnLYVALALALAspGLyGLuArgCysProHisANVal 54
 QY 416 GCGTGGGTCAGAGT-----GGAGAAATGCAAGTAGTAACAGAGCTCGAAGATCC-- 470
 DB 55 GLYTPR-METProGLYPhenAsnGLuArgThrArgGLNArgGLYAspPheProGLYThrTY 74
 QY 471 ----GAGTTC-----CACTGCTGAAGCATGGCCA-----TCAATGA 505
 DB 74 rVALGluPheLeuGLYProVALALALeuALALArgProGLYProALArgProALArg 94
 QY 506 AATTATACGAGCAGCATATCCAGCATCCGAATACCAATACCACTTATCCAGCA 565
 DB 94 gPROLeuProALALArgProLeuAspGLYSerSerGLYSerGLYSerGLYSerGLY 113
 QY 566 TATGATGAG-----GCTCCCTTCTGTACTCTCC----- 596
 DB 113 PLeuALGLuGlnPheSerProProAspProALArgProALArgLeuVALYSLeuVALGL 133
 QY 597 -----AAGCTTCTGTCTCCACTGAGAGACAAACAGG----- 629
 DB 133 uALALLeuGlnALGLuLeuAspSerGLYSerGLYSerGLYSerGLYSerGLYSerGLY 153
 QY 630 -----GATGTAGGATGAGACAGAGAGTAGAGAGAGAGAGAGAGAGAGAGAGAG 682
 DB 153 rArgThrAspThrAspThrAspThrAspThrAspThrAspThrAspThrAspThrAsp 173
 QY 683 CAGAAAGCCAAAGCTTTAAAGATTCAATACACAAATAAACAAGACTCTCTTGCACAC 742
 DB 173 a-----VALYSGLYPheLeu-----LeuALALe 181
 QY 743 TCCTGGCGCTCTATTCATCTCTCCCAAGAGTACCAACCTTACCCAGCAGCAGCA 802
 DB 181 uProALALALALValValThr-----ProGLUAl 190
 QY 803 GGAGAGCAGTGCATATCTGCTCCAAAGCCACCTTCCAGAGAGTCCAGAGAGGAGCCAG 862
 DB 190 aALALALGLuLALArgLALArgLALArgLALArgLALArgLALArgLALArgLALArg 203
 QY 863 GCAGAGAGAGTGCAGAAAGACTTCACTAGAGTCTTGGAGCAGAGAGAGAGAGAGAG 922
 DB 204 -----ProVALGLYLeuVALLeuGLUProProThLeuProLeuH 217

QY 923 GACAAAGCCAGATCTT-----CTTGCCCATCATCAACCAAAACACAGAGAGTCC 976
 DB 217 IseGLNAlALeuThrLeuALArgPheLeuGLNHISLeuGLYArgVALALALArgALALe- 236
 QY 977 ACCTGCCATTGGCAGAGTCTTCTACATGCCAGGAAGACACATATGAACGAGAGACCA 1036
 DB 237 -----ProSerProAspThrALALValALHISALALeUALSerLALPheGLYProLeuLeu 255
 QY 1037 TACAGGTACATGACAC-----ACAG 1057
 DB 255 euArgLeuProProSerGLYGLYGLYAspGLYSerGLUProVALProAspPheProV 275
 QY 1058 TCCTGCTCAGAGATATGCCAAGCTGCAGCCAGCCAC-----AGCCCTCG 1099
 DB 275 ALLeuLeuLeuGLuLALArgLeuVALGLN--GLNHISVALGLuGLuAspALALALProPr 294
 QY 1100 AATGCTGCCCTATGAAACACAAACACTCGGAGAACCTGACCCACCA----- 1145
 DB 294 cALALeuPro---ProLYSProSerLYSALALALYSProALALProThrALALeUALAsnGL 313
 QY 1146 -----AAGCTGATGAGAAAGATGCTGCAGAAATGATGATGATGATGATGATGATGAT 1198
 DB 313 YGLYSerProProSerLeuGLNAspALA-----GLuThrPYrTYrGLYAspPheLe 330
 QY 1199 TCGCCAGCAGCAGTGAAGATGTGTATATGAAGACAAAGATGCTATTTTGTGCTCG 1258
 DB 330 rArgGLuGLuVALAsnGLuLALArgLeu--ArgAspThrProAspGLYThrPheLeuVALr 349
 QY 1259 AGACTGCTCTACAAATTCAGAGCAGACACATATGTTGGTGGGTTTATGGGAGACAA 1318
 DB 349 gAspLALSerSerLYSLIEGLNGLYGLU---TYrThrLeuThrLeuALGLYSGLYGLYAs 368
 QY 1319 GGTCTACATGTGAATATCCGTTCTCTCGAGACCAATCAACAGATTTGGCTGGCAGCAG 1378
 DB 368 n-----AsnLYSLeuLeuLYSVALrPheHISArgAspLYSLYrGLYPheserGLUPr 386
 QY 1379 ACTACGAGGAATGATGATTTGATTTGTTGAGAGACATCTTGAACATCACTACATNT 1436
 DB 386 oLeu-----ThrPheCYSerVALGLuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 401
 RESULT 13
 P85A_BOVIN
 ID P85A_BOVIN STANDARD; PRT; 724 AA.
 AC P23727;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
 p85-alpha subunit) (PtdIns-3-kinase p85-alpha) (PI3K).
 GN PI3KRI.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 OX NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9191567; PubMed=1707345;
 RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Larrea F., Panayotou G.,
 RA Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,
 RA Courtnedge S.A., Parker P.J., Waterfield M.D.;
 RT Characterization of two 85 kd proteins that associate with receptor
 RT tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase.";
 RL Cell 65:91-104(1991).
 RN [2]
 RP CIRCULAR DICHROISM AND FLUORESCENCE SPECTROSCOPY.
 RX MEDLINE=93049176; PubMed=1330535;
 RA Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski B., Dhand R.,
 RA Fry M.J., Blundell T.L., Mollmer A., Waterfield M.D.;
 RT Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
 RT domain with a PDGF receptor phosphorylation site: structural features
 RT and analysis of conformational changes.";
 RL EMBO J. 11:4261-4272(1992).

```

RN [3]
RP STRUCTURE BY NMR OF 1-84.
RX MEDLINE=93272320; PubMed=7684655;
RA Booker G.W., Gout I., Downing A.K., Driscoll P.C., Boyd J.,
RT "Solution structure and ligand-binding site of the SH3 domain of the
RT p85 alpha subunit of phosphatidylinositol 3-kinase.";
RL Cell 73:813-822(1993).
RN [4]
RP STRUCTURE BY NMR OF 314-431.
RX MEDLINE=92357146; PubMed=1323062;
RA Booker G.W., Breeze A.L., Downing A.K., Panayotou G., Gout I.,
RT "Structure of an SH2 domain of the p85 alpha subunit of
RT phosphatidylinositol-3-OH kinase.";
RL Nature 358:684-687(1992).
RN [5]
RP STRUCTURE BY NMR OF 321-434.
RX MEDLINE=97110350; PubMed=8952511;
RA Guenther U.L., Liu Y., Sanford D., Bachovchin W.W., Schaffhausen B.;
RT "NMR analysis of interactions of a phosphatidylinositol 3'-kinase SH2
RT domain with phosphotyrosine peptides reveals interdependence of major
RT binding sites ";
RL Biochemistry 35:15570-15581(1996).
RN [6]
RP STRUCTURE BY NMR OF 614-724.
RX MEDLINE=98173872; PubMed=9512716;
RA Siegal G., Davis B., Kristensen S.M., Sankar A., Linacre J.,
RA Stein R.C., Panayotou G., Waterfield M.D., Driscoll P.C.;
RT "Solution structure of the C-terminal SH2 domain of the p85 alpha
RT regulatory subunit of phosphoinositide 3-kinase.";
RL J. Mol. Biol. 276:461-478(1998).
CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE PI10 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
CC -1- SUBUNIT: HETERODIMER OF A PI10 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-slb.ch/announce/
CC or send an email to license@isb-slb.ch).
CC -----
DR EMBL; M61745; AAA79511.1; .
DR PIR; A38749; A38749.
DR PDB; 2PNA; 31-JAN-94.
DR PDB; 2PNB; 31-JAN-94.
DR PDB; 2PNI; 31-OCT-93.
DR PDB; 1PNT; 31-OCT-93.
DR PDB; 1BFI; 25-FEB-98.
DR PDB; 1BFJ; 25-FEB-98.
DR InterPro; IPR001720; PI3kinase_P85.
DR InterPro; IPR000198; RhogAP.
DR InterPro; IPR001452; SH3.
DR pfam; PF00017; SH2; 2.
DR pfam; PF00018; SH3; 1.
DR pfam; PF00620; RhogAP; 1.
DR PRINTS; PR00678; PI3KINASEP85.
DR PRINTS; PR00401; SH3DOMAIN.
DR ProDom; PD000093; SH2; 1.
DR ProDom; PD000093; SH2; 2.
DR SMART; SM00324; RhogAP; 1.
DR SMART; SM00325; SH2; 2.
DR SMART; SM00326; SH3; 1.

```

[illegible]

Alignment Scores:

Pred. No.:	0.0911	Length:	3421
Score:	135.00	Matches:	76
Percent Similarity:	30.23%	Conservative:	28
Best Local Similarity:	22.09%	Mismatches:	139
Query Match:	4.42%	Indels:	101
DB:	1	Gaps:	15

US-09-856-061-1 (1-1721) x TRGU_HSVB (1-3421)

```

QY 87 ACTCTGCTGAGGAGAGACATGTCACACTATCTTACAGAGTCTCCAGGATCGACC 146
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2347 ThrLeuAlaGluAspGluIleAlaGluLeuProTyrLeuAsnAlaAspSerLeu 2366
QY 147 GTGAGCCCCCTTCCAGGAGCTAGCCGTCTCAACACTGAGCCCTGACTAAAGAAACT 206
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2367 LeuProPromeTasn-----ProAspAsp----- 2374
QY 207 GAGCAGGCTGAGCTGAGATCCCTCTCTTTGCGCAGGACCAAGCATTGACAGCCAG 266
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2375 -----ProIleTyr-----ThrArgValIleSerGlyThr 2384
QY 267 GGCATTAAGAGACACAGAAAGAGA-----TTCCGTGATCTGAGATTCCAGAACGTC 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2385 AsnIleProThrAlaThrThrGluGlySerLeuPheAlaAsp-----GlnGlnLeu 2401
QY 321 TCTCTGTGAAAAATAGTGCATGCGCAAGCCTCAGCAGTGCACAAAGGGGGTGTGAGCG 380
    |||::|||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2402 GluPheLeuArgProGluSerAsnProPhe----- 2411
QY 381 GTTCTGGAACCACTTCGCGATCACAAGAAAGAACTTGCT--GGGGTCCAGGTGAGAGA 437
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2412 -----ProPheAlaSerHisAspSerSerGlnSerLeuAspValProSerSerPro 2428
QY 438 AATGCAACAGTAACAAGACGACTACGAGATCCT----- 470
    |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2429 SerSerGlySerAspLysTyrGlnGluAspProThrGlyIleValTyrAspAlaProVal 2448
QY 471 -----GAGTTCACCTGCTGAAGCATGCGCATCA----- 500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2449 AspAspMetSerAspMetAlaMetAsnLysAlaLysAlaTrpGlnGluTrpLeuGluAsp 2468
QY 501 -----ATGAAATTTTACCAAGCCGACCTATCCAG 530
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2469 GlyPheAlaGluAspAspTyrArgGluLeuSerAsnAlaMetProAlaProPro---Lys 2487
QY 531 GAATCGAATACGACGATACAGCGTATTTCCAGATATGAGAGCTCCCTCTCTGTTA 590
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2488 ThrThrProValValGluSerLysGlnLysSerHisSerValAspArgAlaProThrLeu 2507
QY 591 CCTCCCAAGGCTTCTGCTCCACTGAGACAAACCAAGGATGTGAGATGACACAGCTG 650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2508 ProProLysAlaAlaProLeuProProSerAspAlaSerAlaIleMetSer----- 2524
QY 651 GAAAGAGTGACAGACCTTACCTTCAAGATGTCAAGAGCCACCGCTTAAAGATCAAA 710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2525 -----GlyLysProVal-----PheLys 2530
QY 711 TACACAAAATAAACAAGACTCTTTGGACACTCTCGGCTGCTATACACTCTCCCAAG 770
    ||||| |||||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2531 TyrThrProGlyAsnLysSerAlaValProProSerValProAla-----ProPro 2547
QY 771 AAGTACCAACCTTACCCCGACACACAGAGAGAGACAGTGCATCTTGCCTCAAG 830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2548 ThrLeuProProAlaProProLeuProGlnSerThrSerLysAlaAlaSerGlyProPro 2567
QY 831 CCCACCTTTCCAGAACTCCAGAGGGGGCCCAAGCAGAGAGTGCAAAAGACTTCAGTAGG 890
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2568 ProThrLeuProProAlaProProLeuProGlnSerThrSerLys----- 2582
QY 891 GTCTTGAGAGAGAAAGAAATCTCACACACAGCAAGCAAGAAATCTTCTTGCCCATCA 950
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2583 -----AlaAlaSerGlyProProProThrLeuProProAlaProProLeuProGln 2599

```

```

QY 951 TCAAAACAAAACACACAGAAAGATCCACCTGCCA---TTGCCAGCTTCTTCTACATGCCA 1007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2600 SerThrSerLysAlaAlaSerGlyProPro-ProThrLeuProProAlaProProLeuPr 2619
QY 1008 GGAAGCACCA 1017
    ::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 2619 OGInSerThr 2622

```

Search completed: April 21, 2003, 12:22:26
 Job time : 50.7582 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:15:31 : Search time 112.318 seconds
(without alignments)
6314.347 Million cell updates/sec

Title: US-09-856-061-1

Perfect score: 3057
Sequence: 1 acgagggccaactgcacagc.....taaaaaaaaaaaaaaaaaa 1721

Scoring table:

BLOSUM62
Xgapop 10.0, Ygapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 671580 segs, 206047115 residues

Total number of hits satisfying chosen parameters: 1343160

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+*n2p*.model -DEV-xlp
-Q/cgn2.1/USPTO.spool/US09856061/runat_21042003_113019_3744/app-query.fasta_1.3150
-DB=SPTREMBL_21 -QFMT=fastlan -SUFFIX=n2p.rspt -MINMATCH=0.1 -LOOPCL=0
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061.eccn1_1.324.etrnat_21042003_113019_3744 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGIO -DEV_TIMEDOUT=120
-NARN_TIMEDOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: SPTREMBL_21:*
2: sp.archaea:*
3: sp.bacteria:*
4: sp.fungi:*
5: sp.human:*
6: sp.invertebrate:*
7: sp.mammal:*
8: sp.mhc:*
9: sp.organella:*
10: sp.phage:*
11: sp.plant:*
12: sp.rodent:*
13: sp.virus:*
14: sp.vertebrate:*
15: sp.unclassified:*
16: sp.virus:*
17: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2316	75.8	435	11	Q9JMJ3 mus musculi

2	2305	75.4	435	11	Q9QZE2	Q9QZE2 mus musculi
3	1132	37.0	376	4	Q9P2U9	Q9P2U9 homo sapien
4	410.5	13.4	533	11	Q922M0	Q922M0 mus musculi
5	410	13.4	525	11	Q920L0	Q920L0 rattus norv
6	383	12.5	530	13	Q9DC07	Q9DC07 gallus gall
7	319	10.4	456	4	Q8WV28	Q8WV28 homo sapien
8	315.5	10.3	456	4	Q75498	Q75498 homo sapien
9	309.5	10.1	433	4	Q75499	Q75499 gallus gall
10	297.5	9.7	457	11	Q9QUN3	Q9QUN3 mus musculi
11	287.5	9.4	457	11	Q88504	Q88504 mus musculi
12	287.5	7.2	297	11	Q9D413	Q9D413 mus musculi
13	219.5	6.5	516	11	Q54737	Q54737 mus musculi
14	200	5.5	516	11	Q9UCX5	Q9UCX5 homo sapien
15	167.5	5.5	594	4	Q92529	Q92529 homo sapien
16	165.5	5.4	594	4	Q92529	Q92529 homo sapien
17	165.5	5.4	594	4	Q8TAP2	Q8TAP2 homo sapien
18	156.5	5.1	812	15	Q85466	Q85466 y73 sarcoma
19	154.5	5.1	217	11	Q9CX99	Q9CX99 mus musculi
20	152	5.0	486	4	Q9UCX4	Q9UCX4 homo sapien
21	151.5	5.0	2089	4	Q14676	Q14676 homo sapien
22	149.5	4.9	572	4	Q9UG31	Q9UG31 homo sapien
23	149.5	4.9	645	4	Q95928	Q95928 homo sapien
24	149.5	4.9	1254	4	Q96A74	Q96A74 homo sapien
25	149.5	4.9	1263	4	Q9UPV8	Q9UPV8 homo sapien
26	148.5	4.9	1029	12	Q9YMX0	Q9YMX0 lymantria d
27	147	4.8	2090	4	Q96OC2	Q96OC2 homo sapien
28	143.5	4.7	662	5	Q9N355	Q9N355 caenorhabdi
29	142.5	4.7	1400	4	Q9UP87	Q9UP87 homo sapien
30	141.5	4.6	474	11	Q61120	Q61120 mus musculi
31	141	4.6	728	4	Q9UPH9	Q9UPH9 homo sapien
32	139.5	4.6	286	5	Q95PX0	Q95PX0 caenorhabdi
33	139.5	4.6	398	5	Q95PW9	Q95PW9 caenorhabdi
34	139	4.5	17352	5	Q95YM2	Q95YM2 procambarus
35	138.5	4.5	335	4	Q96C28	Q96C28 homo sapien
36	138.5	4.5	926	4	Q9HOK1	Q9HOK1 homo sapien
37	138	4.5	559	11	Q91252	Q91252 mus musculi
38	138	4.5	655	11	Q8R122	Q8R122 mus musculi
39	138	4.5	535	11	Q91991	Q91991 caenorhabdi
40	137.5	4.5	594	11	Q70143	Q70143 rattus norv
41	137	4.5	477	11	Q70142	Q70142 rattus norv
42	136.5	4.5	1360	16	Q98BD7	Q98BD7 rhlzobium l
43	135.5	4.4	498	13	Q93291	Q93291 figu rubrip
44	135	4.4	854	11	Q9QZ58	Q9QZ58 mus musculi
45	134.5	4.4	309	11	Q9QV20	Q9QV20 mus sp. shb

ALIGNMENTS

RESULT 1
ID Q9JMJ3 PRELIMINARY: PRT; 435 AA.
AC Q9JMJ3;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE MIST.
GN CLNK OR MIST.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20208989; Pubmed=10744659;
RA Goltisuka R., Kanazashi H., Sasamune H., Fujimura Y., Hidaka Y.,
RA Tatsuno A., Ra C., Hayashi K., Kitamura D.;
RT "A BASH/SLP-76-related adaptor protein MIST/Clk involved in Ige
RT receptor-mediated mast cell degranulation.";
RL Int. Immunol. 12:573-580(2000).
DR EMBL; AB021220; BAA96240.1; -;
DR HSSP; P23727; 2PNB.
DR MGD; MGI:1351468; Clnk.
DR InterPro; IPR000980; SH2.

DR Pfam: PF00017; SH2; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS00001; SH2; 1.
 DR SEQUENCE 435 AA: 49513 MW: 401BD3E2FC61ED6 CRC64;

Alignment Scores:

Pred. No.: 1,29e-197 Length: 435
 Score: 2316.00 Matches: 435
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 75.76% Indels: 0
 DB: 11 Gaps: 0

US-09-856-061-1 (1-1721) x Q9JMU3 (1-435)

QY 255 ATGACACACGAGGCAATATAAGCAACGAAGAAGATTGCGTATCGATTCGATTCGAC 314
 DB 1 MetThrSerGlnGlyAsnLysArgThrThrLysGlnGlyAspLeuAArgPheGln 20
 QY 315 AACGTCTCTGCTGTAATAATAGTCATGCGCAAGCCTCAGACAGTGCAGAGGGGGGTGT 374
 DB 21 AsnValSerLeuLeuLysAsnArgSerTrpProSerLeuSerSerAlaLysGlyArgGly 40
 QY 375 CGAGCGGTTCTGGAACCACTCCGGATCACAAGAAAGAACTTGCGTGGGTCACAGTGA 434
 DB 41 ArgAlaValLeuGlnProLeuProAspHisArgArgAsnLeuAlaGlyValProGlyGly 60
 QY 435 GAAAAATCAACAGTAAACAGACAGTACAGATCTGATGCTCCAGCTCTCAAGGCTG 494
 DB 61 GlnLysCysAsnSerAsnAsnAspTyrGlnLysProGlnProGlnLeuLysAlaLysTrp 80
 QY 495 CCATCAATGAAATTTTACACGACGACGACCTATCCAGAAATGGGAATACGAGATACAGC 554
 DB 81 ProSerMetLysIleLeuProAlaArgProIleGlnLysGlnLysAlaAspThrArg 100
 QY 555 TATTTCCAGGATATGATGAGAGGCTCCCTTGTATACCTCCCAAGGCTTGTCTGCACT 614
 DB 101 TyrPheGlnAspMetMetGlnAlaProLeuLeuLeuProLysAlaSerValSerThr 120
 QY 615 GAGAGACAACACGAGATGTGAGATGACAGCTGGAGAAAGTGAAGCAAGCCCTACTTC 674
 DB 121 GlnArgGlnThrArgAspValArgMetThrGlnLeuGlnValAspLysProThrPhe 140
 QY 675 AAGATGTACAGAACCAACGCTTTAAAGATTCAATATACAAAAATTAACAAGACTCTCT 734
 DB 141 LysAspValArgSerGlnArgPheLysGlyPheLysTyrThrLysIleAsnLysThrPro 160
 QY 735 TTGGCACCCTCTCGCTGCTATCATCTCTCCCAAGAAATACCAACCTTACCCCGACGA 794
 DB 161 LeuProProAlaArgProAlaIleThrLeuProLysLysTyrGlnProLeuProAla 180
 QY 795 CCACACAGAGAGAGACATGATCTGCTCCCAAGCCCAACCTTCCAGAAAGTCCAGAG 854
 DB 181 ProProGlnLysSerSerAlaTyrPheAlaProLysProThrPheProGlnValGlnArg 200
 QY 855 GGGCCCAAGCAGAGAGATGCAAAAGACTTCAAGAGGCTCTTGAGAGCAGAAGAAATCT 914
 DB 201 GlyProArgGlnArgSerAlaLysAspPheSerArgValLeuGlyAlaGlnLysLysSer 220
 QY 915 CACACACGACAAAGCCAGAAATCTTCTGCGCATCATCAACCAAAACACAGAGAGAT 974
 DB 221 HisHisGlnThrLysProGlnLysSerCysProSerSerSerAsnGlnAsnThrGlnLysSer 240
 QY 975 CCACTGCGCATTCGACGCTTCTCTACATGCAAGAAAGACAGTATACAAAGCCAGAGAC 1034
 DB 241 ProProAlaIleAlaLysSerSerTyrMetProGlyLysHisSerIleGlnAlaAsp 260
 QY 1035 CATACAGATGATGATGACAGACTCTCTGCTCAGAGATGCCAAGCTCAGCCACGACAGC 1094
 DB 261 HisThrGlySerMetGlnHisCysProAlaGlnArgCysGlnAlaAlaLysSerHisSer 280

QY 1095 CCTCGAATGCTGCTGCTATGTAACCAACAACTCGGAGAAAGCTGACCCACAAAGCTGAT 1154
 DB 281 ProArgMetLeuProTyrGlnAsnThrAsnSerGlnLysProAspProThrLysProAsp 300
 QY 1155 GAGAAGATGCTGTCACAAATGAATGTGATACATTTGGAGAAATGACGCGCCAGCAGTGA 1214
 DB 301 GlnLysAspValIlePheLysGlnLysGlnLysGlnLysGlnLysGlnLysGlnLys 320
 QY 1215 GATGTGTTAATGAAGACAGACAGATGTAATTTTGGTCCAGAGCTCTACAAAA 1274
 DB 321 AspValLeuMetLysGlnAsnLysAspGlyThrPheLeuValArgAspCysSerThrLys 340
 QY 1275 TCACAGCAGACCAATATGTTTGGTGTGTTTATGGGAACAGGCTCTACAAATGTA 1334
 DB 341 SerLysAlaGlnProTyrValLeuValAlaPheTyrGlnLysValTyrAsnValLys 360
 QY 1335 ATCCGTTTCTCGAGACAGCAATCAACAGTTTCCCTGGGACAGAGATCAGGAATGAG 1394
 DB 361 IleArgPheLeuGlnLysSerAsnGlnGlnPheAlaLeuGlnLysLeuArgGlyAsnGln 380
 QY 1395 ATGTTTGATTTCTGTGGAAGACATCATTTGACACTACACATATTTTCCATTTCTAATA 1454
 DB 381 MetPheAspSerValGlnAspIleIleGlnHisTyrThrTyrPheProIleLeuLeuIle 400
 QY 1455 GATGGAAAGACAGAGCTGACGAGCAAGAAACAGTGTACCTCACCAGCAGCTGCTCTC 1514
 DB 401 AspLysLysAspLysAlaAlaArgArgLysGlnCysTyrLeuThrGlnProLeuProLeu 420
 QY 1515 GCCAGGCTCTTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1559
 DB 421 AlaArgLeuLeuLeuThrGlnTyrSerSerGlnAlaLeuHisGln 435

RESULT 2

Q9QZE2 PRELIMINARY: PRT: 435 AA.
 AC Q9QZE2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE CLINK.
 GN CLINK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6;
 RX MEDLINE=20029811; PubMed=10562326;
 RA Cao M.Y., Davidson D., Yu J., Latour S., Veilleux A.;
 RT Clink, a Novel SLP-76-related Adaptor Molecule Expressed in Cytokine-
 RT Stimulated Hemopoietic Cells.";
 RL J. Exp. Med. 190:1527-1534(1999).
 DR EMBL: AF187819; AAF14299.1; -;
 DR HSSP: P23727; 2PNB.
 DR MGI:1351468; Clink.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS00001; SH2; 1.
 DR SEQUENCE 435 AA: 49492 MW: 50D27EC971FC0B5 CRC64;

Alignment Scores:

Pred. No.: 1,23e-196 Length: 435
 Score: 2305.00 Matches: 433
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 75.40% Indels: 0
 DB: 11 Gaps: 0

US-09-856-061-1 (1-1721) x Q9QZE2 (1-435)


```

Db 61 ThTrpInSerIleLeuProAlaArgProIleLysIleuSerIuTyAlaAsp 80
Qy 549 ACAGCTATTTCAGGATATGATGAGGCGCTCTGTACTCCCAAGCTTCTGTC 608
Db 81 ThHisTyrPheLysValAlaMetAspThrProLeuProLysPThrIgtHisSerIle 100
Qy 609 TCCAGTGAAGACAAACAGGAGTGTAGATGACACACTGGAGAGATGACAGACCT 668
Db 101 SerIleGlyIleProThrIleThrIlePheSerIleGln---ThiArgLeuGlnArgValAspIlePro 119
Qy 669 ACCTTAAGAGTGTCAAGAGCCCAAGCTTTAAAGATTCAATACCAAAATAAACAAG 728
Db 120 IleSerArgPheValArgSerGlnAsnIleLysGlyAspAlaSerValArgLysAsnLys 139
Qy 729 ACTCTTTTCCACCTCTGCTGCTGCTATCAGTCTCCCAAGAGATGACACCTTACCC 788
Db 140 IleProLeuProProArgProLeuIleThrIleProLysIleTyIleProLeuPro 159
Qy 789 CCAAGCACCAAGAGAGAGAGATGATCTGCTCCCAAGAGCCCACTTTCCAGAGATC 848
Db 160 Pro---GluProGluSerSerArgProProLeuSerGlnArgHisThrPheProGluVal 178
Qy 849 CAGAGGGGGCCAGGAGAGAGTGCAGAAAGACTTCAGTAGGCTCTTGAGACAGAGAA 908
Db 179 GlnGlyMetProSerIleIleSerLeuArgAspLeuSerGlnValIleuGlnAlaGlnLys 198
Qy 909 GAATCTCACACCAAGACCAAGACAGATCTTCTGCGCCATCATCAACCAACCAACAG 968
Db 199 ValProHisAsnGlnArgIleProGluSerThrHisLeuGlnAlaAsnGlnAsnThrGln 218
Qy 969 AAGAGTCCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1028
Db 219 GlnIleProLeuAlaIleSerSerSerPheThrHisSerValHisSerValGlnAsn 238
Qy 1029 AAGAGCATATACAGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 1088
Db 239 ArgAspHisArgGlyGlyMetIleProCysSerProGlnArgCysGlnProProAlaSer 258
Qy 1089 CACAGCCCT---CGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1142
Db 259 CysSerProHisGlnAlaAsnIleLeuProTyIleHisSerIleThrArgProProHePro 278
Qy 1143 ACAAGCCCTGATGAGAGATGCTGCGAGATGATGATGATGATGATGATGATGATGATG 1202
Db 279 LysArgSerArgPheValAspValGlnHisAsnGlnIleGlyIleGlyIleGlySerArg 298
Qy 1203 CAGGAGTGAAGATGCTGTTATGAGAAAGAGAGATGATGATGATGATGATGATGATG 1262
Db 299 GlnAlaValGlnGlnAlaPheMetLysGlnAsnLysAspGlySerPheLeuValArgAsp 318
Qy 1263 TGCCTTACAAATCCAGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1322
Db 319 CysSerThrLysSerLysGlnGlnProTyValIleuAlaValPheTyGlnAlaAsnVal 338
Qy 1323 TACATGTGAAGATCCGTTCTGCGAGAGATGAGATGAGATGAGATGAGATGAGATG 1382
Db 339 TyrAsnValLysIleArgPheLeuGlnArgAsnGlnGlnPheAlaLeuGlyThrGlyLeu 358
Qy 1383 CGAGAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1430
Db 359 ArgGlyAspGlnLysPheAspSerValGlnLysPheIleIleGlnHisIleTyI 374

```

RESULT 4

```

0922M0 PRELIMINARY: PRT: 533 AA.
AC 0922M0;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
CN Lymphocyte cytosolic protein 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (Apr-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC006948; AA006948.1; -.
DR MGD: MGI:1321402; Lcp2.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR PROSITE: PS0001; SH2; 1.
SQ SEQUENCE 533 AA: 60238 MW: 50AEAD025EP0AD01 CRC64:

Alignment Scores:
Pred. No.: 8,896-28 Length: 533
Score: 410.50 Matches: 148
Percent Similarity: 39.68% Conserved: 50
Best Local Similarity: 29.66% Mismatches: 176
Query Match: 13.43% Indels: 125
Db: Gaps: 20

US-09-856-061-1 (1-1721) x 0922M0 (1-533)
Qy 294 TTCGGTGAATGAGATTCAGACGTCCTCTGCTG-----AAAATAGC 338
Db 57 PheProLysLeuArgMetProLeuLeuSerLysLeuSerGlnAspIleAsnLysAsnGln 76
Qy 339 TCATGGCCAGGCTGAGCAGCAGCAGGCGGTGTCGAGCGGTCTGGAACCACTTCG 398
Db 77 GluArgArgSerIlePheThrArgLysProGlnIleProArgPheLeuGlnIleThrL 96
Qy 399 GATCACAGAGCACTTGCTGCTGCG-----422
Db 97 SerHisGlnGlnAspAspGlyGlyTyrSerSerPheGlnAspArgTyGluSerProAsn 116
Qy 423 -----GTCCAGGTGGAGAAATGCAACAGTAAACAGCACTGAGAGATCTT----- 470
Db 117 AspAspAspProAspGlyGln-----AspAspGlyAspTyGluSerProAsnGln 133
Qy 471 GAGTTCACGCTGCTG-----AAGCATGGCCATCA-----500
Db 134 GlnGlnGlnAlaLeuValAspAlaAlaAspTyGluProProProSerAsnAsnGln 153
Qy 501 -----ATGAATTTTACCAAGCCAGCACTATCCAG-----GAATCGGAATAC 542
Db 154 GlnAlaLeuGlnSerSerIleLeuProProAsnSerPheHisAsnThrAsnSerMetTy 173
Qy 543 GCAGATACACGCTATTTCCAGAGATGATGAGAGGCTCCCTTCTGTAACTT-----CCC 596
Db 174 IleAspArgProProThrGlyLysValSerGlnIleProProValProProLeuArgPro 193
Qy 597 AAGCTTCTGTC-----608
Db 194 LysProAlaLeuProProLeuProThrGlyArgAsnHisSerProLeuSerProHis 213
Qy 608 -----608
Db 214 ProAsnHisGlnGluProSerArgSerGlyAsnAsnLysThrAlaLysLeuProAlaPro 233
Qy 609 TCCAGTGAAGACAAACAGGAGATGAGATGACACAGCTGGAAGAGTGCAG-----662
Db 234 SerIleAspArgSerThrLysProProLeuAspArgSerLeuAlaProLeuAspArgGln 253
Qy 663 -----AAGCTTACTTCAAGATGTCAGAAAGCCCAAGCTTTAAAGGA 704
Db 254 ProPheIleLeuGlyLysLysProProPheSerAspLysProSerAlaProLeuGlyArg 273
Qy 705 TTCAAATACAAATAAACAAGACTCTTTCAGCACTCTGCGCTGCTATACACTCTC 764
Db 274 GlnHisLeuProLysIleGlnLysProProLeuPro-----ProAlaMetAspArg 290
Qy 765 CCCAAGAGATACCAACCTTTACCCCGACAGCACCAAGAGAGAGAGATGATGATGCTGCT 824

```


OY	1206	GCAGTGTGATGGTTGTTATGAAGAAGACAAAGATGCTACTTTTTGGTCGAGACTGC	1265
Dd	423	GUUAGLUAIALALALEuAArgLLysIleasnInAspLIyThrPheueValArAspSer	442
OY	1266	TCTACAATAATCCAGGCAAGCACATATGTTTGCGTGTATTATCGAACAAAGCTAC	1322
Dd	443	SerLySThrValAsnAsnProTy-ValLeuMetValLeuTyrlYsAspLyValTyr	462
OY	1326	AATGTGMAAATCCGCTTCTCGAAGCAATCACAATTGCCCTGGGCGACAGACTACGA	1385
Dd	463	AsnIleGlnIleArgTyrlGlnGluIsrGclValTyrlLeuLeuLylThrylGlyLeuArg	482
OY	1386	GGAATGTGATGTTGATTTCTGTGAGACATCATTTGACATACACACATATTTCCATT	1445
Dd	483	GlyLySGluAspPheLeuSerValSerAspIlelleAspTyrlPheArgLyMetProLeu	502
OY	1446	CCTGTAATATGATGGAAAGACAGCGCTGCACGAGAAACAGTCACTCAACC	1499
Dd	503	LeuLeuIleAspGIyLyLSasnarGlYser---ArgTyrlGlnCythrIleuThr	519
RESULT	6		
OY	09DG07	PRELIMINARY:	PRT; 530 AA.
AC	09DG07;		
DT	01-MAR-2001 (TREMBLrel, 16, Created)		
DF	01-MAR-2001 (TREMBLrel, 16, Last sequence update)		
DT	01-MAR-2002 (TREMBLrel, 20, Last annotation update)		
.DE	SLP-76 adaptor protein.		
GN	SLP76.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Archosauiria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;		
OX	NCBITaxID=9031;		
OK	[1]		
RN	SEQUENCE FROM N.A.		
RP	TISSUE=THYMUS;		
RC	MEDLINE=20451093; PubMed=10939315;		
RX	Liphal M., Kurosaki M., Inabe K., Chan A.C., Sugamura K., Kurosaki T.;		
RA	"Involvement of LAT, Gads, and Grb2 in Complementatation of SLP-76 to		
RT	the Plasma Membrane."		
RL	J. Exp. Med. 192:847-856(2000) .		
DR	EMBL; AF226988; AAC18493.1; -		
DR	HSSP; P12931; ISHD.		
DR	Interpro: IPR001660; SAM.		
DR	Interpro: IPR000980; SH2.		
DR	Pfam; PF00017; SH2; 1.		
DR	PRINTS; PR00401; SH2DOMAIN.		
DR	ProDom; PD000093; SH2; 1.		
DR	SMART; SM00454; SAM; 1.		
DR	SMART; SM00252; SH2; 1.		
DR	PROSITE; PS50001; SH2; 1		
SO	SEQUENCE 530 AA; 59580 MW; CFCBCBCEBD79909A CRC64;		
Alignment Scores:			
Pred. NO.:	2,51e-25	Length:	530
Score:	383.00	Matches:	137
Percent Similarity:	42.34%	Conservative:	62
Best Local Similarity:	29.15%	Mismatches:	166
Query Match:	12,53%	Indels:	105
DB:	13	Gaps:	20
US-09-856-061-1 (1-1721) x 09DG07 (1-530)			
OY	197	AAGGAGACTGACGACGCTGATTAAGATCCCTCTCTTTGCCAGGTGCCAAGACAT	256
Dd	122	LySGLIuAspGLuAlaAsp---TyrluSerProThrgluGluThrgluGlnAlaGlnAs	140
OY	257	GACCACGACGG-----CAATAAAGACAACAAGAGAGATTGCGTAGCTAG	307
Dd	141	AspSerAspGIyLyScIuProProProSerAsnAsnsp-GluAlaIstAsnValII	160
OY	308	ATTTCGAACAGCTCTCTCTGTAATAAATAGTCAATG-----CCAAGCTCAGCAG	358

Db	160	ephr0ialLylsSerLeuAlaSmnshtrAspTyrIleAspArgProPthrSerAr	180
Qy	359	TGCCAAAGGCGGTGTGCAGCGGTTCTGGACACCTTCGGATGACAGAAAGACTGGC	418
Db	180	gSerSerHisGln-----ProValProProGlnArgProGlyProse	195
Qy	419	TGGGGTCCCA-----GGTGGAGAAAAATGCAACAGTAAACAGACTACGAAGATCC	469
Db	195	rProAlaProIaSerPheIleYglArgGlyAlaSer-----	207
Qy	470	TGAATTCACCTCTCGAAGCAGTGGCCATCATGAATTTATTCACAGCAGACCTATCCA	529
Db	208	-----LeuProAlaPheProIeUleuProIylSmnSmnSparGlnValY	223
Qy	530	CGAATGGGAATACGCGAGATACACGCTATTTCAGAGATGTAGTGAGGCTCCCTTGTT	589
Db	223	sProSerLys-----	226
Qy	590	ACCTCCCAAGGCTTGTCTCCATCTAGAGACAAACCGAGATGTAGATGACACAGCT	649
Db	227	-ProProAlaProSerIle-----AspIleSerThrLys-----ProProIa	240
Qy	650	GGAAGAAGTGGACAAACCTACCTTCAGAGATGTCAAGACCAACGCTTTAAAGATTCAA	709
Db	240	uAspArgLeuGlyProPro-----	246
Qy	710	ATACACAAATAATTAACAGACTCT-----TGCC	739
Db	247	-PheGlnArgGlnAsnProValProGlyArgLysProGlyHisProGlnLysLeuLeuTh	266
Qy	740	ACCTCTGGCGCTGTATCACTCTCCCAAGATACACACCTTGACCCACAGCACC	799
Db	266	rProGlnLeuArgAlaLeuGlyGlnGlnLeuAlaMetIleProLysProValProPr	286
Qy	800	AGAGGAGCAGCATGCATCTTCGCTCCAAAGCCACCTTTCCAGAA-----GTCCAGAG	853
Db	286	oSerAspArgTyrGlnArgGlyAsnProSerProLeuArgGlnIleProValYsG1	306
Qy	854	GGGGCCCAAGCAGAGAGAGTGCMAAAGACTTCAGTAGAGGTCCTTGAGCAGAGAGAGATC	913
Db	306	ncGlyTrpAlaGlnGlnLysArgProGln-----GlnGlnGlnLys	319
Qy	914	TCAC-----CACCGACAAAGCCAGAAATCTTCTGGCCATCTCAACCAAAACACA--	965
Db	319	pHisIleProGlnArgAlaValProGlnIleSerLeuProProTyrIleSerSerAsnThrPh	339
Qy	966	-----CAGAGATCCACCTGCATTCGACGCTCTCTTACATCCAGGAAA	1012
Db	339	eProSerLysSerIleLysAlaProProLysProGlySerAsnSer-----IleProGlyAl	358
Qy	1013	GCAGAGTATACAGCCAGACAGACCACTACAGTAGACATGACAGCACTGCTGCTCAGAGATG	1072
Db	358	agLysIleAlaArgSerLeuSerAlaSerGlySerLeu-----ProProArgPhePr	375
Qy	1073	CCAACTGCACCCAGCAGACACCCCTCGAATG-----CTGCC	1108
Db	375	oLeuLysSmnSerIleArgSerProSerArgGlyThrIleAlaSerLeuArgProIlePr	395
Qy	1109	CTATCAAAACACAAACTGGAGAAACCTGACCCACAAAGCCTGATGAGAAAGATGCTCG	1168
Db	395	oIleProSerArgIleThrIleAlaHisGlnThrAsnThrGlnGlnAspGlu-----AspSerIe	414
Qy	1169	GCAGAAATGAAGGATACGTGGAGATATACAGTCCCGCAGCAGTGGAGAAAGTGTATTATGA	1228
Db	414	uAsnSpsolunThrPyrValAlaTyrIleSerArgProGlnAlaGlnAlaIleAlaLeuArgL	434
Qy	1229	AGAGAACAGAGATGTACTTTTGTGTCGAGACTGCTTACAAATGCAAGCCAGAAC	1288
Db	434	sIleAsnLysAspLylThrPheLeuValArgAspSerSerArgLysThrIleThrHisPr	454
Qy	1289	ATATGTTTTGTGTGTTTTATGGGAACAAGCTTACAAATGTGAATAATCCGTTTCTCGA	1348

Db	454	cttvttaileumetvalleu	tyllylsaspryvaly	tyrasnllleglnlleatgtytgcgl	474
Qy	1349	GAGCAATCAACAGATTGGCCCTGGGACAGACTACAGAGAAATGAGATGTTGATTCGT			1408
Db	474	uglnasprglnthrtyleuendclgthrcglyleuylsglylssguasphseser	val		494
Qy	1409	GGAAGACATCATTCAGACACATCAACATATTTGCCATTCGTCAATAGATGGGAAGACAA			1466
Db	494	lalanstlellelaspryrtgheglalnrgthrtproleuleuileasglylsaspar			514
Qy	1469	GGCTGCACGCAGGAACAGTGTACCTC			1496
Db	514	ggllyser--ArgAsnGlnGlyValLeu			522
RESULT 7					
08MW28	ID	08MW28	PRELIMINARY:	PRT:	456 AA.
AC	08MW28;				
DT	01-MAR-2002 (TEMBLrel. 20, Created)				
DT	01-MAR-2002 (TEMBLrel. 20, Last sequence update)				
DT	01-JUN-2002 (TEMBLrel. 21, Last annotation update)				
DE	B-cell linker.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=TONSIL;				
RA	Strausberg R.;				
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.				
DR	EMBL; BC018906; AAH18906.1; -				
DR	InterPro: IPR000980; SH2.				
DR	Pfam: PF00017; SH2; 1.				
DR	SMART: SM00252; SH2; 1.				
DR	PROSITE: PSS0001; SH2; 1.				
SO	SEQUENCE 456 AA; 50465 MW; ED6D424A035D1792 CRC64;				

RT Interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
 RT receptor activation.";
 RL J. Biol. Chem. 272:27362-27368(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=98361304; PubMed=9697839;
 RA Fu C., Turk C.W., Kurosaki T., Chan A.C.;
 RT "BLNK: a central linker protein in B cell activation.";
 RL Immunity 9:93-103(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20050956; PubMed=10583958;
 RA Minglish Y., Rohrer J., Coustan-Smith E., Lederman H.M., Pappu R.,
 RA Campana D., Chan A.C., Conley M.E.;
 RT "An essential role for BLNK in human B cell development.";
 RL Science 286:1954-1957(1999).
 DR EMBL: AF068180; AAC39936.1;
 DR EMBL: AF180756; AAF20382.1;
 DR EMBL: AF180740; AAF20382.1; JOINED.
 DR EMBL: AF180741; AAF20382.1; JOINED.
 DR EMBL: AF180742; AAF20382.1; JOINED.
 DR EMBL: AF180743; AAF20382.1; JOINED.
 DR EMBL: AF180744; AAF20382.1; JOINED.
 DR EMBL: AF180745; AAF20382.1; JOINED.
 DR EMBL: AF180746; AAF20382.1; JOINED.
 DR EMBL: AF180747; AAF20382.1; JOINED.
 DR EMBL: AF180748; AAF20382.1; JOINED.
 DR EMBL: AF180749; AAF20382.1; JOINED.
 DR EMBL: AF180750; AAF20382.1; JOINED.
 DR EMBL: AF180751; AAF20382.1; JOINED.
 DR EMBL: AF180752; AAF20382.1; JOINED.
 DR EMBL: AF180753; AAF20382.1; JOINED.
 DR EMBL: AF180754; AAF20382.1; JOINED.
 DR EMBL: AF180755; AAF20382.1; JOINED.
 DR HSSP: P29354; 1BM8.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 SQ SEQUENCE 456 AA; 50466 MW; 95FID5485D03D397 CRC64;

Alignment Scores:
 Pred. No.: 1 21e-19 Length: 456
 Score: 319.00 Matches: 99
 Percent Similarity: 39.65% Conservative: 58
 Best Local Similarity: 25.00% Mismatches: 139
 Query Match: 10.44% Indels: 100
 DB: 4 Gaps: 12

US-09-856-061-1 (1-1721) x 075498 (1-456)

QY 444 AACGTAACAAGACGACTAGAGAGATTCGAGTCCAGCTGAGGAGCATGCGCATCATG 503
 DB 91 AsnAlaAspAspSerTYRGLuProProProValGlu-----GlnGluThrArg 106

QY 504 AAAATTTTACAGCCAGCACTATCCAGATCGAATACGACATACACGCTATTTCCAG 563
 DB 107 ProValHisProAlaLeuProPheAlaArgGlyGluTYRLeaAspAsnArgSerSerGln 126

QY 564 GATATGATGAGCGCTCC-----CTTCTGTACTCCCAAGCGTCTGTCTTCACAT 614
 DB 127 -----ArgHisSerProProPheSerLysThrLeuProSerLysProSerThrProSer 144

QY 615 GAGAGACAAACAGGAGTGATGAGATGACACAGCTGAGAGAGTGGACAGCCTACCTTC 674
 DB 145 GluTysAlaArgLeuThrSerThrLeuProAlaLeuThrAlaLeuGlnLysProGlnVal 164

QY 675 AAGGATTCAGAACCCACACGCTTTAAAGATTCAATACACAAATAAACAAGACTCCT 734
 DB 164 ----- 164

QY 735 TTGCACACTCTCGGCTGCTATCATCTCCCAAGAAAGTACCAACCTTACCCCGACGA 794

DB 165 -----ProProLysProLysGlyLeuLeuGlnAspGlnAlaAspTYRValValProVal 182

QY 795 CCACCAGAGAGACAGATGCATCTTCGCTCCAAAGCCACCTTCCAGAAAGTCCAGAG 854

DB 183 GluAspAsnAspGluAsnTYRLeuHisProThrGluSerSerSerProProGluLys 202

QY 855 GGGCCCGAGGAGAGAGTCAAAAGACCTTCAGTAGGTCTTGGAGCAGAAAGAAATCT 914

DB 203 AlaProMet-----Val 206

QY 915 CACCCACAGCAAAAGCAGAAATCTTCTGCCATCATCAAAACCAACACA----- 965

DB 207 AsnArgSerThrLysProAsnSerThrProHisSerProProGlyThrAlaSerGly 226

QY 966 -----CAGAAAGTCCACCTGCTGACAGCTTCCTACATG 1004

DB 227 ArgAsnSerGlyAlaTrpGluThrLysSerProPro-----AlaAlaProSerProLeu 245

QY 1005 CCA-----GGAAGCAC-----AGTATACAAAGCAGAGAC 1034

DB 246 ProArgAlaGlyLysLysProThrThrProLeuLysThrThrProValAlaSerGln 265

QY 1035 CATACAGGTACCATG-----CAGCAGTCTCTGCTCAGAGATGCCAGCTGCA--- 1082

DB 266 AsnAlaSerSerValCysGluGluLysProLeuProAlaGluAlaGlnHisArgLysSer 285

QY 1083 -----GCCAGCCACAGCCCTCGAATGCTGCCATATGAAACACAACTCGGAG 1130

DB 286 HisArgGlnGluAlaValGlnSerProValPheProProAlaGlnLysGlnLeuHisGln 305

QY 1131 AAACCTGACCCC----- 1142

DB 306 LysProLeuProLeuProAlaArgPheThrGluGlyLysAsnProThrValAspGlyProLeu 325

QY 1143 -----ACAAAGCCTGATGAGAAGCATGCTGCGACAGATGAA 1178

DB 326 ProSerPheSerSerAsnSerThrHisSerGluGlnGluAlaGlyValLeuCysLysPro 345

QY 1179 TGTACATTTGGAGAAATACAGTCCGACGACGAGTGAACATGTGTTAATGAAGACACAG 1238

DB 346 TrpTYRAlaGlyAlaCysAspArgLysSerAlaGlnLeuAlaLeuHisArgSerAsnLys 365

QY 1239 GATGCTACTTTTGTGTCGAGACTGCTTACAAATATCCAAAGCAGCAACCATATGTTTG 1298

DB 366 AspGlySerPheLeuLeuArgLysSerSerGlyHisAspSerLysGlnProLYRThrLeu 385

QY 1299 GTGCTGTTTATGGCAACAAGTCTACATGTGAATAATCCGTTTCTCGAGACCAATCAA 1358

DB 386 ValValPhePheAsnLysArgValTYRAsnIleProValArgPheLeuGluAlaThrLys 405

QY 1359 CAGTTTGCCCTGGGCGACAGACTACGAGAAATGAGATGTTTATCTGTGGAAGCATC 1418

DB 406 GlnTYRAlaLeuGlnArgLysLysAsnGlyGluGluTYRThrGlySerValAlaGluIle 425

QY 1419 ATTGAACATACACATATTTTCCCATCTGCTATATAGATGGAAAGAC 1466

DB 426 IleArgAsnHisGlnHisSerProLeuValLeuIleAspSerGlnAsn 441

RESULT 9

09YGC1 PRELIMINARY; PRT; 552 AA.

ID 09YGC1

AC 09YGC1

DT 01-MAY-1999 (TREMBLER, 10, Created)

DT 01-MAY-1999 (TREMBLER, 10, Last sequence update)

DT 01-MAR-2002 (TREMBLER, 20, Last annotation update)

DE B cell linker protein BLNK.

GN BLNK OR BASH.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9916381; PubMed=10023776;
 RA Ishii M., Kurosaki M., Pappu R., Okawa K., Ronko I., Fu C.,
 RA Shibata M., Iwamatsu A., Chan A.C., Kurosaki T.;
 RT "BLNK required for coupling syk to PLC gamma 2 and Rac1-JNK in B
 RT cells.";
 RL Immunity 10:117-125(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=99049801; PubMed=9834055;
 RA Gotsuka R., Fujimura Y., Mamada H., Morimura T.,
 RA Uetsuka K., Doi K., Tsuji S., Kitamura D.;
 RT "BASH, a novel signaling molecule preferentially expressed in B cells
 RT of the bursa of Fabricius.";
 RL J. Immunol. 161:5804-5808(1998).
 DR EMBL; AF089727; AAD12783.1; -;
 DR EMBL; AB015289; BAA36275.1; -;
 DR HSSP; P08487; 2PLD.
 DR InterPro: IPR002965; P_rich_extensn.
 DR InterPro: IPR000980; SH2.
 DR Pfam; PF00017; SH2; 1.
 DR PRINTS; PRO1217; PRICHEXTENS.
 DR PRINTS; PRO0401; SH2DOMAIN.
 DR PRODOM; PD000093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR PROSITE; PS50001; SH2; 1.
 DR SEQUENCE 552 AA; 61822 MW; FB32179BE38D072 CRC64;

Alignment Scores:

Pred. No.:	2,63e-19	Length:	552
Score:	315.50	Matches:	115
Percent Similarity:	32.58%	Conservative:	43
Best Local Similarity:	23.71%	Mismatches:	138
Query Match:	10.32%	Indels:	189
DB:	13	Gaps:	16

US-09-856-061-1 (1-1721) x 09YGC1 (1-552)

QY	423	GTCCAGAGTGGAGAAATTCAGACAGTACAGACACTGCAAGATCTGAGTCCAGCTG	482
DB	105	ValProSerGluGlu-----AsnProAspSerTyrGluProProProSerGluGln	122
QY	483	CTGAAGGCGATGGCCATCATGAAATTTTACAGCCAGCAGATATCCAGAAATCGGAATAC	542
DB	123	GluLys-----LysLysIleProSerSerPheProIleSerArgGlyGluTyr	138
QY	543	GCAGATACACGCTATTTCCAGATATGATGAGGCTCCCTTCTGTTACTCCCAAGGCT	602
DB	139	AlaAspAsnArgThrSerHisIleGln-----LeuProPro-----	150
QY	603	TCTGTCTCCACGTAGAGACAAACAGGAGTGTGAGGATGACACACCTGGAGAAGATGAC	662
DB	151	-----IleAsn	152
QY	663	AAGCTTACCTTCAGAGATGTCAAGCAACGCTTTAAAGATCAAAATACACAAATA	722
DB	153	LysPro-----Leu	155
QY	723	AACAAGACTCT-----TTGCCA-----CCTCTCGGCTGCTATCATCTCCCAAG	770
DB	156	ProSerThrProSerSerAlaLeuProArgProLysProSerLeuProSerProAla	175
QY	771	AAGTACCAACCTTACCCCCACACCA-----	797
DB	176	AlaLysProLysLeuProLeuLysProArgGlyLysSerAspAspGlnAspAsnTyrIle	195
QY	798	-----CCAGAGAGAGAGAGATGATCTTC	821
DB	196	ValProValAspAsnAspAspAspAsnTyrIleGluProThrGluSerSer-----	212
QY	822	GCTCCAAAGCCACCTTTCACAGAGTCCAGAG-----	854

DB	213	ThrProProAlaLysProProValAsnArgPheMetLysProProAlaLysSerAla	232
QY	855	-----GGGCCAGAGGAGAGAGTGCAGAAAGACTTCAGTGGTCTTGGACAGAA	905
DB	233	LeuProThrProProLysProSerLeuAlaSerAspMetGlnGluValTyrGluValPro	252
QY	906	GAAGAATCTCCACCACAG-----ACAAAGCCAGAACTTCTP	941
DB	253	GluGluGluGluGluLeuSerProProProValThrArgPheThrLysProLeuProAla	272
QY	942	TGCCCATCATCAACCAAAAC-----ACACAGAGATCCACCT	980
DB	273	ThrArgAlaLysAlaGlnIleSerHisMetHisSerMetThrArgGluSerProLys	292
QY	981	GCATTCGCAGCTCTTCTCATGATGCA-----GGAAG	1013
DB	293	LeuAspAlaSerArgAsnIleLeuProLeuProArgAsnArgLeuHisProLysThrAsp	312
QY	1014	CACAGTATACAGCCAGACAGACATACAGTAGACATGCAGCAGCAGCTGCTCAGAGATGC	1073
DB	313	HisGluAlaAsnAsnAsnAsnArgLysAsnHisSerPheSerAsnThrGlnGluSerPhe	332
QY	1074	CAGCTCAGCCAGCCAGCCAGCT-----	1097
DB	333	ProProGluAlaAlaProSerProLeuProArgAlaLeuLysLysThrSerAsnAlaVal	352
QY	1098	-----CGAATGCTGCCCTATGAAACCA-----AACTGGAGAAACT	1136
DB	353	AsnProAlaLysProCysLeuProSerArgAspThrPheThrValAsnGluAspLysPro	372
QY	1136	-----	1136
DB	373	ThrAlaAlaAspArgArgArgGlySerSerHisGluPheProLeuProProLeuProSer	392
QY	1136	-----	1136
DB	393	GlyThrProLysSerSerLeuGlnLysProLeuValLeuProLysValProGluAlaPro	412
QY	1137	-----GACCCCAAAAGCT	1151
DB	413	SerArgAlaLeuGlyThrSerProHisSerSerIleSerSerIleSerSerThrAlaAsp	432
QY	1152	GATGAGAAGAGTGTGCGCAGATGATGATGATGATGATGATGATGATGATGATGATGATG	1211
DB	433	GlnAspArgLysValHisSerLysAlaArgLysAlaArgLysAlaArgLysThrAla	452
QY	1212	GAAATGTGTATTAAGAAGAAACAGATGCTCTTTTGGTGGCGAGCTGCTACA	1271
DB	453	GluAspAlaLeuTyrArgSerAsnLysAspLysSerPheLeuIleArgLysSerSerGly	472
QY	1272	AAATCCAGGAGCAACCATATGTTTGGTGTGTTTATGGCAACAGCTGCTACATGTG	1331
DB	473	GlnAspSerArgGlnProTyrThrLeuValAlaPheTyrAsnArgArgValTyrAsnIle	492
QY	1332	AAATCCGTTTCTCGAGACAGATCAACATGATGCTTGGCTGGCAGCAGCTGAGGAAT	1391
DB	493	ProIleArgPheIleGluSerThrArgGlnTyrAlaLeuGlyArgGlyLysCysGlyLeu	512
QY	1392	GAGATGTTGATTCGTGGAGACATCATGTGACCTACATATTTCCATCTGCTGTA	1451
DB	513	GluArgPheAspSerValAlaGluIleValGluAsnHisGlnHisThrSerLeuValLeu	532
QY	1452	ATGATGGGAAAGAC	1466
DB	533	IleAspSerGlnAsn	537
RESULT 10			
ID	075499	PRELIMINARY;	PRT; 433 AA.
AC	075499;		
DT	01-NOV-1998 (TREMBlrel. 08, Created)		
DT	01-NOV-1998 (TREMBlrel. 08, Last sequence update)		
DT	01-MAR-2002 (TREMBlrel. 20, Last annotation update)		


```

Db 428 AsnSerHisGlnHisAsnProLeuValLeuIleAspSerGlnAsn 442
RESULT 12
088504
ID 088504 PRELIMINARY: PRT; 457 AA.
AC 088504;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE B cell linker protein BLNK.
GN LY57
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001722; PubMed=9341187;
RA Fu C., Chan A.C.;
RT "Identification of two tyrosine phosphoproteins, pp70 and pp68, which
RT interact with phospholipase Cgamma, Grb2, and Vav after B cell antigen
RT receptor activation.";
RL J. Biol. Chem. 272:27362-27368(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98361304; PubMed=9697839;
RA Fu C., Turck C.W., Kurosaki T., Chan A.C.;
RT "BLNK: A Central Linker Protein in B Cell Activation.";
RL Immunity 9:93-103(1998).
DR EMBL: AF068182; AAC40206.1; -.
DR HSSP: P23727; IBFT.
DR MGD: MGI:96878; LY57.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS5001; SH2; 1..
SQ SEQUENCE 457 AA; 50803 MW; 66D235796A6C45F0 CRC64;

Alignment Scores:
Pred. No.: 7,79e-17 Length: 457
Score: 287.50 Matches: 106
Percent Similarity: 38.31% Conservatlive: 53
Best Local Similarity: 25.54% Mismatches: 139
Query Match: 9,40% Indels: 117
Db: 11 Gaps: 16

US-09-856-061-1 (1-1721) x 088504 (1-457)
QY 450 AACAGCACTGAGAGATCCGAGTCCAGCTGCGAAGCATGGCCATCAATGAAATT 509
Db 69 AspSerAspTyrGlnAspProAspGlnHis-----SerAspSerGlnMetTyrVal 85
QY 510 TTACAGCAGCAACATATCCAGCAATCGAATACGAGATACAGCATATTCAGAGATATG 569
Db 86 MetProAlaGluGluThrGlyAspAspSerTyrGlu----- 97
QY 570 ATGAGAGCTCCCTCTGTACTCTCCAGGCTTCTGTCTCCAGTGAAGAGCAAAACGAG 629
Db 98 -----ProProProAla-----GluGlnGlnThrArg 106
QY 630 GATGTAGAG-----ATGACACAGCTGAGAGAGATGAGACAGCCATCCTTCAAG 677
Db 107 ValValHisProAlaLeuProPheThrArgGlyGluTyrValAspAsnArgSer----- 124
QY 678 GATGTAGAGCAACCAAGCTTTAA-----GATTCAAATATACAAAATAAACAAGACT--- 731
Db 125 -----SerGlnArgHisSerProProPheSerTyrThrLeuProSerLysProSer 141
QY 732 ---CCTTGGCACTCCTCGGCTGCTATACACTCTCC-----AAGAG 773
Db 142 TrpProSerAlaLysAlaArgLeuAlaSerThrLeuProAlaProAsnSerLeuGlnLys 161

```

```

QY 774 TACCAACCTTACCCCGACACACCA-----GAGAGAGCAGTGCATCTTCGCT 824
Db 162 ProGln---ValProProLysProLysAspLeuLeuGlnAspGlnAspTyrValVal 180
QY 825 CCAAAACCCACCTTCCAGAAAGTCCAGAGGGGGCCAGCAGAGAGT----- 872
Db 181 ProValGlnAspAsnAspGlnAsnTyrTleHisProValGluSerSerProProAla 200
QY 873 -----GCAAAAGACTTCAGAGGCTCTTGAGAGCAGAGAAGAAATCTCACCAGCACA 926
Db 201 GlutylAlaProMetValAsnArgSerThrLysProAsnSerSerLysHisMetSer 220
QY 927 AAGCCAGATCT----- 938
Db 221 ProProGlyThrValAlaGlyArgAsnSerGlyValTrpAspSerLysSerLeuPro 240
QY 939 TCTTGCCCATCATCAACCAAAACACAGAGAGTCCA----- 977
Db 241 AlaAlaProSerProLeuProArgAlaGlyLysLysProAlaThrProLeuLysThrThr 260
QY 978 -----CCTGCCATTGCCAGCTCTTCTCATCATGCCAGAAACACAGATACACCCAGA 1031
Db 261 ProValProProLeuProAsnAlaSerAsnValLysGluGlnLysProValProAlaGlu 280
QY 1032 GACCATACAGTAGCATGCAGCAGCTGCTCTGCTCAGAGATGCCAAGTGCACGCCAGC 1091
Db 281 ArgHisArgLysSer-----SerHis 287
QY 1092 -----AGCCCTGAAATGCTGCCCTTGAACACAAACATCGAGAA 1133
Db 288 ArgGlnAspThrValGlnSerProValPheProThrGlnLysProValHisGlnLys 307
QY 1134 CCTGACCCCA----- 1145
Db 308 ProValProLeuProArgPheProGlnAlaGlySerProAlaAlaAspGlyProPheHis 327
QY 1146 -----AAGCTGATGAGAAAGATGTCGTGGCACAATGAATG 1181
Db 328 SerPheProPheAsnLeuThrPheAlaAspGlnGluGlnLeuGlyLysProTrp 347
QY 1182 TACATTGAGAAATACAGTCCGACGAGCATGGAAGATGTGTAATGAAGACAAAGAT 1241
Db 348 AlaPhePheAsnLysArgValTyrAsnLeuProValArgPheIleGlnAlaThrLysGln 407
QY 1362 TTTCCCTGGGCGACAGACATCAGAGAAATGAGATGTTGATCTGTGAAGACATCAT 1421
Db 408 TyrAlaLeuGlyLysLysLysAsnGlyGluGlnLysPheGlySerValValGluIleVal 427
QY 1422 GAACACTACACATATTTTCCATTCCTGCTAATACATGGGAAAGC 1466
Db 428 AsnSerHisGlnHisAsnProLeuValLeuIleAspSerGlnAsn 442

RESULT 13
Q9D413
ID Q9D413 PRELIMINARY: PRT; 297 AA.
AC Q9D413;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 21, Last annotation update)
DE 4933424C13RIK protein.
GN 4933424C13RIK
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```



```

Db 103 AspSerAspTyrGluAsnProAspGluHis-----SerAspSerGluMetTyrVal 119
QY 510 TTTCAGCCAGACAGCTATCCAGGAATGGAATACGAGATACAGCTATTCCAGGATATG 569
Db 120 MetProAlaGluGluThrGluAspSerTyrGlu----- 131
QY 570 ATGGAGAGCTCCCTGTGTACCTCCAGAGGCTTGTGTCTCCACATGAGACAAACAGG 629
Db 132 -----ProProProAla-----GluGluGlnThrThrArg 140
QY 630 GATGTGAGC-----ATGACACAGCTGGAGAAAGTGGACAAAGCTTACCTCAAG 677
Db 141 ValValHisProAlaLeuProPheThrArgGluTyrValAlaAspTyrValVal 160
QY 678 GATGTGAGAAAGCCAAAGCTTAAAGATTC-----AAATAC 713
Db 161 ArgThrLeuArgProSerLeuAlaArgHisPheProValSerProAlaGluGlnArgLys 180
QY 714 ACAAATAATGAACAGACTCTTTGCCACCTCTCCGCTGATATCACTCTCCCAAGAG 773
Db 181 AlaGluLeuGluAlaSerThrLeuProAlaProAsn-----SerLeuGlnLysProGln--- 198
QY 774 TACCAACCTTACCCSCACACCACCA-----GAGGAGACAGTGCATCTTGCT 824
Db 199 -----ValProTyrLeuTyrLeuAspLeuGluGlnAspGluAlaAspTyrValVal 215
QY 825 CCAAAAGCCACCTTCCAGAGTCCAGAGGCGCCAGGAGAGAGTGCACAAAGACTTC 884
Db 216 ProValGluAspAsnAspGluAsnTyrTleHisProArgGlnSerSerAlaSerArgLeu 235
QY 885 ACTAGGGCTCTTGAAGCAGACAGAAAGATCTCACACCAAAAGCCAAAGATCTTCTGC 944
Db 236 LeuAlaArgLeuProTyrSerThrLeuAspGlnProSerGlnThrValProGlnSerThrCys 255
QY 945 CCATCA-----TCAAAACCAAAACACACAGAAAGTCCACCTCCATTCSCAGCT 995
Db 256 AlaSerAlaArgValAlaGluGlnThrValGluSerGluThrProSerHisLeuGluSer 275
QY 996 TCTCATGATCCAGAAAGCAGATATA-----CAAGCCAGAGACCATACAGTACGATG 1049
Db 276 ProHisHisPheGlnHisGluGlnArgLeuArgSerGlnLeuHisHisLeuArgLeu 295
QY 1050 CAGCACTGTCTCTCAGACAGATGCCA----- 1076
Db 296 ProPheLeuProTyrGluMetHisGlnMetPheValLysLysSerLeuPheLeuLeuSer 315
QY 1077 -----GCTGACAGCCAGCCAC-----AGCCCTGATGCTGCCCTAT 1112
Db 316 AlaAsnArgGlySerSerHisArgGlnAspThrValGlnSerProValPheProThr 335
QY 1113 GAAAACAAACCTCGAGAAACCTGACCCACAA----- 1145
Db 336 GlnLysProValHisGlnLysProValProLeuProLysAlaGluSerProAlaHisAsp 355
QY 1146 -----AACCTGATGAAAGATGCTGCG 1169
Db 356 GlyProPheHisSerPheProPheAsnSerThrPheAlaAspGlnGluAlaGluLeuLeu 375
QY 1170 CAGAAATGAATGATACATGGAAATACAGTCCAGCAGCAGTGAAGATGTTAATGAA 1229
Db 376 GlyLysProTyrTyrValAlaProValAsnArgLysSerAlaGluGlnAlaLeuHisArg 395
QY 1230 GAGAACAAAGATGCTATTTTGGTCCGAGACTGCTCTACAAATCCAAAGCAGAACCA 1289
Db 396 SerAsnLysAspGlySerPheLeuLeuArgLysSerSerGlnHisAspSerLysGlnPro 415
QY 1290 TATGTTTGGTGGTGTATTAAGGAGAACAGTACATGTGAATCCGTTCTCTCGAG 1349
Db 416 TyrThrLeuValAlaPhePheLysSerLacysIleIlePheLeuTyrGluLeuLeuLys 435
QY 1350 AGCAATCAAGATTTGCCCTGGGCAAGGACTACGAGGAATGAGATTTG----- 1401
Db 436 GlnPro-AsnSerMetLeuTyr-----GluArgArgLysMetValLysSe 450

```

```

QY 1402 -----ATTCTGTGAGACATCATTAACACTACACATATTTCCATCTGCTA 1451
Db 450 rThrSerGluValLeuThrGlySerSerThrValIleSerThrTrn--ProLeuValLeu 469
QY 1452 ATAGATGGAAAGACAAAGCTGCACGC 1478
Db 470 IleAspSerGlnAsnAsnSerLysArg 478

RESULT 15
Q9UCX5 PRELIMINARY: PRT: 471 AA.
ID Q9UCX5
AC Q9UCX5:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE RAI, SHC-51.9 kDa SHC-related protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96330343; PubMed=8760305;
RA Pellicci G., Dente L., De Giuseppe A., Verducci-Galletti B., Gluli S.,
RA Mele S., Vetranti C., Giorgio M., Pandolfi P.P., Cesaroni G.,
RA Pellicci P.G.,
RT "A family of Shc related proteins with conserved PTB, CH1 and SH2
RT regions."
RL Oncogene 13:633-641(1996).
DR HSSP: P29353; ISHC.
DR InterPro: IPR000050; PTD_domain.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00640; PTD; 1.
DR Pfam: PF00017; SH2; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00629; SHCPTDOMAIN.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00462; PTB; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS01179; PTD; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50001; SH2; 1.
SO SEQUENCE 471 AA; 51969 MW; 90521E77D1589CAD CRC64;

Alignment Scores:
Pred. No.: 3,88e-06 Length: 471
Score: 167.50 Matches: 106
Percent Similarity: 36.40% Conservative: 60
Best Local Similarity: 23.25% Mismatches: 161
Query Match: 5.48% Indels: 129
DB: 4 Gaps: 25

US-09-856-061-1 (1-1721) x Q9UCX5 (1-471)
QY 291 GGATTCGGTATCTGAGATCTCAGAACGCTCTCTGTGAATAATAGTCAATGCGCAAGC 350
Db 91 GlySerSerAsnLeuGlnPheAlaGluMetSerThrLeuThrIleSerThrAlaSer 110
QY 351 CTCAGCACTGCCAAAGGGCGGTGCGAGCGTCTGAAACCACTTCGGATACAGAAAG 410
Db 111 LeuAsnLeuArgThrProAspSerLysGlnIleIle-----AlaAsnHisMetArg 128
QY 411 AACCTGCTGGGCTCCAGAGTGGAGAAATAATGCAACAGTAAACAGACTC----- 461
Db 129 SerIleSerPheAlaSerGlyGlyAspPro-----AspThrThrAspTyrValAlaTyr 146
QY 462 -----GAAGATCT-----GAGTTCAGCTGCTGAAGGCGATGCGCATCAATG 503
Db 147 ValAlaLysAspProValAsnArgAlaCysHisIleLeuGlnLysCysAspGly--- 165
QY 504 AAAATTTTACACGCAACACCTATTCACAGAAATGGAATACGACAGATACACCTATTTCAG 563
Db 166 -----LeuAlaGlnAspValIleGlySerThrIleGlyGlnAlaPheGlnLeuArgPheLys 183

```


QY 564 GATATGATGAGGCTCCCTTCTGTACCTCCCAAGCTTCTGTCTCACTGAGAGACA 623
Db 184 GlnTyrLeuGlnCysProThrLysIlePro-----AlaLeuHisAspArgMetGln 200
QY 624 ACCAGGATGTGAGTACACAGCTGAGAGAAAGTGGACAGAGCTTCAAGATGTC 683
Db 201 SerLeuAspGlnProThrPheGln---GlnGlnGlnLysPhe----- 212
QY 684 AGAAGCCAAAGCTTAAAGATCAATACACAAAATAAACAAGACTCTTGCACCT 743
Db 213 -----GlySerAspHisProTyrTyrAsnSerIleProSerLysMet 226
QY 744 CCTGGCCTCTATCACTCTCCCAAGAGATACCAACCTTACCCCAAGCCACAGAG 803
Db 227 ProProPheGlnLysPheLeuAspThrArgLeuLysProArgProHisAlaPro----- 244
QY 804 GAGAGCAGTGCATCTTCCGCTCCAAG----- 830
Db 245 ---AspThrAlaGlnPheAlaGlnLysGlnGlnThrTyrTyrGlnGlnArgHisLeuGly 263
QY 831 CCCACCTTCCAGAA---GTCCAGAGGGGGCCAGCAGAGAGTGCACAAAGACTTCAGT 887
Db 264 AspThrPheGlnLysAspThrGlnGlnThrProLeuArgGlnLysSerAsp----- 281
QY 888 AGGCTCTTGGAGCAGAGAGATCTCACACCACAGACAAAGCCAGAGATCTTGGCCA 947
Db 282 -----IleTyrSerThrProGlnGlnLysLeuHisValAlaProThrGlnGlnAlaPro 299
QY 948 TCATCAACCAACACACAGAGAGTCCACCTGCACTTGCAGCTTCTCTACATGCCA 1007
Db 300 ---ThrTyrValAsnThrGlnGlnIleProPro----- 309
QY 1008 GGAAGACAGGTATACAGCCACAGACCAATACAGTACGACACTGCTGCTCAG 1067
Db 310 -----GlnAlaTyrProAla--- 314
QY 1068 AGATGCCAAGCTGAGCCAGCCACAGCCCTCGA-----ATGCTGCCCTAT 1112
Db 315 ---AlaValSerSerAlaGlnSerSerProArgLysAspLeuPheAspMetLysProPhe 333
QY 1113 GAAACACAAACTCGGAGAAACT----- 1136
Db 334 GlnAspAlaLeuLysAsnGlnProLeuGlnProValLeuSerLysAlaAlaSerValGln 353
QY 1137 -----GACCCACAAAGCTGATGAGAG-----GATGCTGGCAG 1172
Db 354 CysIleSerProValSerProArgAlaProAspAlaLysMetLeuGlnGlnLeuGlnAla 373
QY 1173 AATGAATGTACATTGGAGATACAGTGCAGGAGCTGAGAGATGTAAATGAAGAG 1232
Db 374 GlnThrTyrPheGlnGlnGlnMetSerArgLysGlnAlaGlnGlnLeu----- 390
QY 1233 AACAGAGATGTACTTTTGTCCGAGACTGCTTACAAATCCAGGACAGACCA-- 1289
Db 391 GlnLysAspGlnAspPheLeuValArg-----LysSerThrThrAsnProGly 406
QY 1290 ---TATGTTTGGTGGTGTATTATGGAACAAGCTTACAAATCGAAATCCGTTTCTC 1346
Db 407 SerPheValLeuThrGlnMetLysAsnGlnGlnAlaLysHisLeuLeuValAspPro 426
QY 1347 GAGAGCAATCAACAGTTTGGCCTGGACAGACTACAGAGAAATGAGATGTTGATCT 1406
Db 427 Gln-----GlyThrGlnArgThrLysAspArgValPheAspSer 439
QY 1407 GTGAGACATCATTTGACACTAC-----ACATATTTTCCCATCTGCTAATGATGG 1460
Db 440 IleSerHisLeuIleAsnHisHisLeuGlnSerSerLeuProIleVal----- 455
QY 1461 AAGACACAGGCTGCAGCAGAGAAACAGTCTACTCACCCAGCCACTG 1508
Db 456 -----SerAlaGlnSerGlnLeuLys---LeuGlnGlnProVal 467

Search completed: April 21, 2003, 12:28:38
Job time : 130.318 secs

GenCore version 5.1.4 p5.4578
Copyright (c) 1993 - 2003 Comugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:22:47 ; Search time 35.4007 Seconds

(without alignments)
1637.367 Million cell updates/sec

Title: US-09-856-061-2

Perfect score: 2316
Sequence: 1 MTSQGNKRRTEKGFGLRQ.....QPLPLARLLLTQYSSQALHE 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2316	100.0	435	22	AAAB81070
2	1316.5	56.8	428	23	AAE22609
3	1316.5	56.8	443	23	AAE22608
4	1132	48.9	376	22	AAAB81071
5	1109.5	47.9	353	23	AAE22610
6	326	14.1	456	22	AAAB93406
7	319	13.8	449	19	AAAB9866
8	319	13.8	456	19	AAAB9865
9	287.5	12.4	457	19	AAAB9867
10	260	11.2	95	20	AAAB70587

11	213	9.2	84	21	AAAG03152
12	212	9.2	46	22	ABAB8147
13	212	9.2	46	22	ABAB23354
14	212	9.2	46	22	AAAB58783
15	212	9.2	46	22	AAAB71296
16	212	9.2	46	22	AAAB18973
17	212	9.2	46	22	AAAB31575
18	212	9.2	46	22	ABGA41096
19	185.5	8.0	503	23	ABG228554
20	165.5	7.1	474	18	AAAB15253
21	165.5	7.1	594	18	AAAB15256
22	150.5	6.5	743	22	AAAB79738
23	144.5	6.2	474	18	AAAB39089
24	144.5	6.2	474	18	AAAB39087
25	142.5	6.2	474	18	AAAB39084
26	142.5	6.2	474	18	AAAB39092
27	141.5	6.1	474	18	AAAB39085
28	141.5	6.1	474	18	AAAB39088
29	141.5	6.1	474	18	AAAB39090
30	141.5	6.1	474	18	AAAB39091
31	141.5	6.1	474	18	AAAB39095
32	141.5	6.1	474	18	AAAB39096
33	141.5	6.1	474	18	AAAB39097
34	141.5	6.1	474	18	AAAB39098
35	141.5	6.1	474	18	AAAB39099
36	140.5	6.0	474	18	AAAB39094
37	140	6.0	469	17	AAAB97242
38	140	6.0	540	22	AAAB40461
39	140	6.0	640	22	AAAB38675
40	138.5	6.0	474	18	AAAB39086
41	138.5	6.0	474	18	AAAB39093
42	138.5	6.0	923	22	AAAB11773
43	137.5	5.9	594	18	AAAB5631
44	137.5	5.9	594	18	AAAB15260
45	134.5	5.8	787	22	AAAB30364

ALIGNMENTS

RESULT 1
AAB81070
ID AAB81070 standard; Protein: 435 AA.

AC AAB81070;
DT 25-JUN-2001 (first entry)
XX
DE Murine mast cell-specific signal transduction protein.
XX
KW Mast cell; signal transduction; mouse; allergic disease.
XX
OS Mus musculus.
XX
PN JP3146204-B1.
XX
PD 12-MAR-2001.
XX
PE 17-SEP-1999; 99JP-0263778.
XX
PR 17-SEP-1999; 99JP-0263778.
XX
PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
DR WPI, 2001-310022/33.
XX
DR N-PSDB; AAF86139.
XX
PT Mast cell-specific signal-transduction molecule, useful for screening
XX
PT therapeutic compounds for treating allergies, is specifically expressed
XX
PS by mouse mast cell -
XX
Claim 1; Page 7-8; 12pp; Japanese.

CC This sequence represents a murine mast cell-specific signal transduction
 CC protein. The invention includes the cDNA and protein sequences of the
 CC mast cell-specific signal transduction molecule and an expression vector
 CC containing the polynucleotide sequence. The coding sequence of the signal
 CC transduction protein can be used for screening therapeutic compounds
 CC which will be useful for treating allergic diseases.

XX Sequence 435 AA:

Query Match 100.0%; Score 2316; DB 22; Length 435;

Best Local Similarity 100.0%; Pred. No. 4.3e-194; Matches 435; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTSQGNKRTTKEGFDLRFQNVSLKNSRSPSLSSAKGRCRAVLPLPDHRRNLAVPGG 60
 DB 1 MTSQGNKRTTKEGFDLRFQNVSLKNSRSPSLSSAKGRCRAVLPLPDHRRNLAVPGG 60
 QY 61 EKCNNSNDYEDPEFQLLKAMPKMLIPAPRTOESEYADTRFYQDMMEAPLLPPKASVST 120
 DB 61 EKCNNSNDYEDPEFQLLKAMPKMLIPAPRTOESEYADTRFYQDMMEAPLLPPKASVST 120
 QY 121 ERQTRDVMTQLEVDKPTFKDVRSORFKGYTKINKTLPPLPPPAITLPKKYQPLPPA 180
 DB 121 ERQTRDVMTQLEVDKPTFKDVRSORFKGYTKINKTLPPLPPPAITLPKKYQPLPPA 180
 QY 181 PPESSAIFAPKPTPEVQGRGPRORSADFSRYLGAEEESHQTRPESSCPSSNONTOKS 240
 DB 181 PPESSAIFAPKPTPEVQGRGPRORSADFSRYLGAEEESHQTRPESSCPSSNONTOKS 240
 QY 241 PPAISSSYMPGKHSIOARDHTGSMOHCPOARCOAAASHSPMLYEYENTNSKRPPTKPD 300
 DB 241 PPAISSSYMPGKHSIOARDHTGSMOHCPOARCOAAASHSPMLYEYENTNSKRPPTKPD 300
 QY 301 EKDVQWQNEWYIGEYSQAVEDVLAKENKDGTFVLVRCSTKSAEPYVLVFEYGNKYVNVK 360
 DB 301 EKDVQWQNEWYIGEYSQAVEDVLAKENKDGTFVLVRCSTKSAEPYVLVFEYGNKYVNVK 360
 QY 361 IFFLESNQOAFALGTGRGNEMFDSVEDIIEHTYTPILLIDGKDKAARRKOCYLTQPLPL 420
 DB 361 IFFLESNQOAFALGTGRGNEMFDSVEDIIEHTYTPILLIDGKDKAARRKOCYLTQPLPL 420
 QY 421 ARLLTQYSSQALHE 435
 DB 421 ARLLTQYSSQALHE 435

RESULT 2
 AAE22609
 ID AAE22609 standard; Protein: 428 AA.
 AC AAE22609;
 XX
 XX 26-JUL-2002 (first entry)
 DE Human MIST splice variant protein from clone #7.
 XX
 KW Human: mast cell immunoreceptor signal transducer; MIST; immune disorder;
 KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
 KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
 OS Homo sapiens.
 XX
 XX WO200226986-A2.
 PD 04-APR-2002.
 PE 28-SEP-2001; 2001WO-US30593.
 PR 29-SEP-2000; 2000US-237030P.
 XX (BRIM) BRISTOL-MYERS SQUIBB CO.
 XX
 XX Perez-Villar JT, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
 PI

XX WPI: 2002-372126/40.
 DR N-PSDB: AAD35801.
 XX

PT New isolated mast cell immunoreceptor signal transducer polypeptide,
 PT useful for treating immune disorder involving hyperactivity of B- or
 PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
 XX
 XX
 PS Claim 11; Fig 5; 171pp; English.

CC The present invention relates to novel mast cell immunoreceptor signal
 CC transducer (MIST) proteins and polynucleotides encoding such proteins.
 CC MIST sequences of the invention are useful for preventing, treating or
 CC ameliorating a medical condition in mammalian subject. They are useful
 CC for treating an immune disorder involving hyperactivity of B- or T-
 CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
 CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
 CC as targets for therapeutic intervention in immune cell disorders and
 CC inflammatory indications, for diagnosis and/or screening of disorders
 CC or diseases associated with expression of MIST, for screening for
 CC antagonists or inhibitors of the interaction of MIST with cellular
 CC signalling components. They are used in assays that detect activation
 CC or induction of various B and T-cell-related neoplasms or cancers.
 CC Sequences of the invention are also used in gene therapy. The present
 CC sequence is human MIST splice variant protein from clone #7.

XX Sequence 428 AA:

Query Match 56.8%; Score 1316.5; DB 23; Length 428;

Best Local Similarity 62.1%; Pred. No. 1.2e-106; Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

QY 1 MTSQGNKRTTKEGFDLRFQNVSLKNSRSPSLSSAKGRCRAVLPLPDHRRNLAVPGG 60
 DB 1 MNRQGNKRTTKEGSDNLFQNFSLPKNSWPRINSATGQYORMKPLDWERNEPAVLDG 60
 QY 61 EKCNNSNDYEDPEFQLLKAMPKMLIPAPRTOESEYADTRFYQDMMEAPLLPPKASVST 120
 DB 61 AKGSHDDDDYDPELMEETWQSIKILPARPIKESYADTHFKVAMQDPLDTRTSISI 120
 QY 121 ERQTRDVMTQLEVDKPTFKDVRSORFKGYTKINKTLPPLPPPAITLPKKYQPLPPA 180
 DB 121 GQPTWNTQ-TRLEKVDKPSKDVRSQNIKGASVKNKIPPLPPPLITLPKKYQPLPP 178
 QY 181 PPESSAIFAPKPTPEVQGRGPRORSADFSRYLGAEEESHQTRPESSCPSSNONTOKS 240
 DB 179 EPSSRPPLSQRHPTPEVQGRGPRORSADFSRYLGAEEESHQTRPESSCPSSNONTOEI 238
 QY 241 PPAISSSYMPGKHSIOARDHTGSMOHCPOARCOAAASHSP--PMLYEYENTNSKRPPTK 298
 DB 239 PLAISSTFTTNSHVSQNRDRGDMQPCSPQRCPPASCSPHENILPYKYYSMRPPPKR 298
 QY 299 PDEKDVQWQNEWYIGEYSQAVEDVLAKENKDGTFVLVRCSTKSAEPYVLVFEYGNKYV 358
 DB 299 SDRDVQWQNEWYIGEYSQAVEDVLAKENKDGTFVLVRCSTKSAEPYVLVFEYGNKYV 358
 QY 359 VKIFFLESNQOAFALGTGRGNEMFDSVEDIIEHTYTPILLIDGKDKAARRKOCYLTQPL 417
 DB 359 VKIFFLESNQOAFALGTGRGNEMFDSVEDIIEHTYTPILLIDGKDKTGVRHKKOCHLTP 418
 QY 418 LPLARLL 425
 DB 419 LPLTRHLL 426

RESULT 3
 AAE22608
 ID AAE22608 standard; Protein: 443 AA.
 AC AAE22608;
 XX
 XX 26-JUL-2002 (first entry)
 DT
 XX

```

DE Human MIST protein #1.
XX
XX Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX
XX Homo sapiens.
OS
FH key Location/Qualifiers
FT Modified-site 84 /label= Tyrosine_phosphorylation_binding_site
FT Modified-site 111 /label= Tyrosine_phosphorylation_binding_site
FT Domain 306..311 /label= SH3_binding_proline-rich_motif
FT Domain 324..407 /label= SH2_domain
FT Domain /label= SH2_domain
XX
XX WO200226986-A2.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US030593.
XX
XX 29-SEP-2000; 2000US-237030P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Perez-Villar J, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX
XX WPI: 2002-372126/40.
XX
XX N-PSDB; AAD55800.
XX
XX New isolated mast cell immunoreceptor signal transducer polypeptide,
XX useful for treating immune disorder involving hyperactivity of B- or
XX T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
XX
XX Claim 11; Fig 3; 171pp; English.
XX
XX The present invention relates to novel mast cell immunoreceptor signal
XX transducer (MIST) proteins and polynucleotides encoding such proteins.
XX MIST sequences of the invention are useful for preventing, treating or
XX ameliorating a medical condition in mammalian subject. They are useful
XX for treating an immune disorder involving hyperactivity of B- or T-
XX lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
XX as targets for therapeutic intervention in immune cell disorders and
XX inflammatory indications, for diagnosis and/or screening of disorders
XX or diseases associated with expression of MIST, for screening for
XX antagonists or inhibitors of the interaction of MIST with cellular
XX signalling components. They are used in assays that detect activation
XX or induction of various B and T-cell-related neoplasms or cancers.
XX Sequences of the invention are also used in gene therapy. The present
XX sequence is human MIST protein.
XX
XX Sequence 443 AA:
SQ
Query Match 56.8%; Score 1316.5; DB 23; Length 443;
Best Local Similarity 62.1%; Pred. No. 1.3e-106;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;
QY 1 MTSQGNKRTKEGFGDLRFQVSLKNSWPSLSAKGRCRAVLEPLPDHRRNLGAGVGG 60
DB 16 MNRQGNKRTKEGSDNLFQNFSLPKNSWRINSATGQYRMNKPFLDMERNRAAVLDG 75
QY 61 EKCSNNDYEDPEFOLKAMPMSKILPAPIQESEYADTRYFODMEAPLLPPKASVST 120
DB 76 AKGSHDDDDYDPELMEETWQSILKLPAPIKSEYADTRYFKVAMDPLPLDFTSTISI 135
QY 121 ERQTRVAMTULEYDKATFTFDVNSQRFKRYKINKTPLPAPPAITLTKKIQPLPA 180
DB 136 GQPTWNTQ-TFLERVDKPKISKDVRSQNIKGASVAKNKPLPPPPPLITLTKKIQPLP- 193

```

```

QY 181 PRESSAYAPKPTPEVQGRPRORSKDESRVLAGAEEESHQFPKSSCPSSNONTOKS 240
DB 194 EPSSSRPPLSQRTTPEVQMPQISLRDLSEVLEAEKPPHNRKPESTHLENQTOEI 253
QY 241 PPAIASSYWPCKHSIQARDHTGSMOCPARCOAAASHSP--RMLPYENTNSEKDDPTK 298
DB 254 PLAISSSFTTSMHNSVQNRHRGGMQPCSPQRQPPASCPHENILPYKTSMRPFPRK 313
QY 299 PDEKDWQNEWYIGETYSROAVEDYLAKENKGTFLVRDCSTKSAEPYLVVYGGKYN 358
DB 314 SDRKDVQHNEMWYIGETYSROAVEFKMKENKGSFLVRDCSTKSEPPYLVAVYENKYN 373
QY 359 VKIRPLESNOQFALGTGLRGNEEMDSVEDIIEHYTFEPILLIDGDKD-AARRKOCYLTOP 417
DB 374 VKIRPLERNOQFALGTGLRGDEKDSVEDIIEHYKNEPILLIDGDKDGYHRKQCHLTOP 433
QY 418 LPLARLL 425
DB 434 LPLRLHL 441
RESULT 4
AAB81071
ID AAB81071 standard; Protein: 376 AA.
XX
XX AAB81071;
XX
XX 25-JUN-2001 (first entry)
XX
XX Human mast cell-specific immunoreceptor signal transducer.
XX
XX DE
XX KW Mast cell; signal transduction; human; allergic disease; MIST;
XX mast cell-specific immunoreceptor signal transducer.
XX
XX OS Homo sapiens.
XX
XX PN JP3146204-B1.
XX
XX PD 12-MAR-2001.
XX
XX PF 17-SEP-1999; 99JP-0263778.
XX
XX PR 17-SEP-1999; 99JP-0263778.
XX
XX PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.
XX
XX DR WPI: 2001-310022/33.
XX
XX N-PSDB; AAF86140.
XX
XX PT Mast cell-specific signal-transduction molecule, useful for screening
XX therapeutic compounds for treating allergies, is specifically expressed
XX by mouse mast cell -
XX
XX PS Example 1; Page 10-11; 12pp; Japanese.
XX
XX This sequence represents human MIST (mast cell-specific immunoreceptor
XX signal transducer). The invention relates to cDNA encoding a murine mast
XX cell-specific signal transduction protein. Included in the invention are
XX cDNA and protein sequences of the mast cell-specific signal transduction
XX molecule and an expression vector containing the polynucleotide sequence.
XX The coding sequence of the signal transduction protein can be used for
XX screening therapeutic compounds which will be useful for treating
XX allergic diseases.
XX
XX Sequence 376 AA:
SQ
Query Match 48.9%; Score 1132; DB 22; Length 376;
Best Local Similarity 60.9%; Pred. No. 1.5e-90;
Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;
QY 19 FQVSLKNSRSPSLSSAKGRCRAVLEPLPDHRRNLGAGVGGKCSNNDYEDPEFOLK 78
DB 1 FQNFSLPKNSRSPRINSATGQYRMNKPFLDMERNFAAVLDGAKGSHDDDDYDPELMEE 60

```

```

QY 79 AWPMSKILPAPRIQSEVADTRYFQDMMEAPLLPPKASVSTERQTRDYRMTOLEVDK 138
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 TMSQIKILPAPRIKESVADTHFKVAMDPPLDPLDRTSISIGQPTWMTQ-TRLEVDK 119
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 139 TKRDVNSQRFKGGKTKYKINKTLPRLPPRAITLPPKKYQPLPPAPPESSAVFAPKPTPEV 198
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 120 ISRDVNSQNIKGDASVYKRNKIPPLPPRLITLPPKKYQPLP-EPSSRPPLSQRHTEPEV 178
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 199 QGPPRRRSKDKFSRYLVGAEESHQTKPESSCPSSNQNTQKSPRALASSYMPGKHSIDA 258
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 179 QGPPRSQISLRDSEVLEAKVPHNQRPSTHLENNQTOEIPLAISSSFTTSMHSVQ 238
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 259 RDHTGSMOHCPORCQAAASHSP--RMLPYENTNSEKPPDPTRKPEKDVQNMENYIGESR 316
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 239 RDHRGGMQPCSPQRCQPPASCPHENILPYKYSWRPPRPKRSRDKVDQHNEMWIGESR 298
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 317 QAVEDVLMKENDGTFVLVDCSTKSKAEPLYLVFYGKRVYVVKIRFLESNOQFALGTG 376
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 299 QAVEEAFMKENKDGSLVLDSCSTKSKKEBPYLVAFVENKYVNVKIRFLERNOQFALGTG 358
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 377 RGNEMFDSVEDITEHY 392
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 359 RGEKFDSEVDITEHY 374
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 5
AAE22610
ID AAE22610 standard; Protein: 353 AA.
AC AAE22610;
XX
XX 26-JUL-2002 (first entry)
DE Human MIST splice variant protein from clone #12.
XX
XX Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX
XX Homo sapiens.
OS
XX
XX W0200226986-A2.
XX
XX 04-APR-2002.
XX
XX 28-SEP-2001; 2001WO-US30593.
XX
XX 29-SEP-2000; 2000US-237030P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
PA
PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX
XX WPI: 2002-372126/40.
XX
XX N-PSDB: AAD35802.
XX
XX New isolated mast cell immunoreceptor signal transducer polypeptide,
XX useful for treating immune disorder involving hyperactivity of B- or
XX T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma -
XX
XX Claim 11; Fig 8; 171pp; English.
XX
XX The present invention relates to novel mast cell immunoreceptor signal
XX transducer (MIST) proteins and polynucleotides encoding such proteins.
XX MIST sequences of the invention are useful for preventing, treating or
XX ameliorating a medical condition in mammalian subject. They are useful
XX for treating an immune disorder involving hyperactivity of B- or T-
XX lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
XX as targets for therapeutic intervention in immune cell disorders and
XX inflammatory indications, for diagnosis and/or screening of disorders
XX or diseases associated with expression of MIST, for screening for

```

```

CC antagonists or inhibitors of the interaction of MIST with cellular
CC signalling components. They are used in assays that detect activation
CC or induction of various B and T-cell-related neoplasms or cancers.
CC Sequences of the invention are also used in gene therapy. The present
CC sequence is human MIST splice variant protein from clone #12.
XX
XX
SQ Sequence 353 AA;
Query Match 47.9%; Score 1109.5; DB 23; Length 353;
Best Local Similarity 64.4%; Pred. No. 1.2e-88;
Matches 226; Conservative 31; Mismatches 89; Indels 5; Gaps 4;
QY 78 KAMPMSKILPAPRIQSEVADTRYFQDMMEAPLLPPKASVSTERQTRDYRMTOLEVDK 137
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 3 TMSQIKILPAPRIKESVADTHFKVAMDPPLDPLDRTSISIGQPTWMTQ-TRLEVDK 61
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 138 PTFKDVRSQRFKGGKTKYKINKTLPRLPPRAITLPPKKYQPLPPAPPESSAVFAPKPTPE 197
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 62 TKRDVNSQNIKGDASVYKRNKIPPLPPRLITLPPKKYQPLP-EPSSRPPLSQRHTEPE 120
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 198 VORGRORSKAKFSRYLVGAEESHQTKPESSCPSSNQNTQKSPRALASSYMPGKHSIQ 257
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 ISRDVNSQNIKGDSEVLEAKVPHNQRPSTHLENNQTOEIPLAISSSFTTSMHSVQ 180
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 258 ARDHTGSMOHCPORCQAAASHSP--RMLPYENTNSEKPPDPTRKPEKDVQNMENYIGESR 315
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 RDHRGGMQPCSPQRCQPPASCPHENILPYKYSWRPPRPKRSRDKVDQHNEMWIGESR 240
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 316 QAVEDVLMKENDGTFVLVDCSTKSKAEPLYLVFYGKRVYVVKIRFLESNOQFALGTG 375
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 QAVEEAFMKENKDGSLVLDSCSTKSKKEBPYLVAFVENKYVNVKIRFLERNOQFALGTG 300
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 376 RGNEMFDSVEDITEHYTFPLILIDGKDK-AARKQCYLQPLPLALL 425
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 RGEKFDSEVDITEHYKNPFLILIDGKDTGVHRKQCHLTQPLPTRLHL 351
   1:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 6
AAB93406
ID AAB93406 standard; Protein: 456 AA.
AC AAB93406;
XX
XX 26-JUN-2001 (first entry)
DE Human protein sequence SEQ ID NO:12602.
XX
XX Human; primer: detection; diagnosis; antisense therapy; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX EP1074617-A2.
XX
XX 07-FEB-2001.
XX
XX 28-JUL-2000; 2000EP-0116126.
XX
XX 29-JUL-1999; 99JP-0248036.
XX
XX 27-AUG-1999; 99JP-0300253.
XX
XX 11-JAN-2000; 2000JP-0118776.
XX
XX 02-MAY-2000; 2000JP-0183767.
XX
XX 09-JUN-2000; 2000JP-0241899.
XX
XX
XX (HELI-) HELIX RES INST.
XX
XX Oka T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX
XX WPI: 2001-318749/34.
XX
XX Primer sets for synthesizing polynucleotides, particularly the 5602
XX full-length cDNAs defined in the specification, and for the detection
XX and/or diagnosis of the abnormality of the proteins encoded by the

```

PT full-length cDNAs -
 XX Claim 8; SEQ ID 12602; 2537bp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of
 CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.
 XX
 XX Sequence 456 AA:
 SQ
 Query Match 14.1%; Score 326; DB 22; Length 456;
 Best Local Similarity 24.6%; Pred. No. 6.6e-20;
 Matches 99; Conservative 50; Mismatches 141; Indels 112; Gaps 9;
 QY 64 NSNDVDEPDEPQILKAMPKILPAPRIQSEVADTYFDQMEAP---LLLPKAVST 120
 DB 91 NADDSYEPPEVE---QETRPVHPALPFAFGELYDNRSSO--RHSPFSKTLSPKSPWPS 144
 QY 121 EROTRDVRMTQLEEVDPKTFKDVRSQRFKGYTKINKTLPPEPAITLPPKYYOPLPPA 180
 DB 145 EKARLTSTLPALATALOKRQY-----PPKPRGLLEADYVVPV 182
 QY 181 PPESSAYFAKPTTFPEVQKPRORSKADFSVLGAEEESHQTKPSSCPSSQNT--- 237
 DB 183 EDNDENYIHPTESSPPPEKAPM-----VNRSTKPMNSPTPASPPTGASG 226
 QY 238 -----QKSPPAIASS-----SYMPGKHSIQARHTQSMQ 266
 DB 227 RNSGAMETKSPPPAAPSPLPRACKKPTTPLKTPPVASQANASVCGEKPIPAERHRGSS 286
 QY 267 HCPAORCOAAASHPRMLPYENTNSEKRPD-----TKPDEK 302
 DB 287 -----KQEAVQSPVFPFPAQKQIHKQPIPLPRTBEGNPTVDGPLPSFSSNSTISQDA 339
 QY 303 DVMQNEWYIGEYSQAVEDVLMKENKDGTLVRCSTKSKAEPYLVVFGNKYVANKIR 362
 DB 340 GVLCKPMPYAGACDRKSAEALHRSNKDGSFLIRKSSGHDSCQPYTLVVFENKRYNIPVR 399
 QY 363 FLESNQFALGTGLRGNEMFDSVEDIIEHTYTPILLIDGD 404
 DB 400 FIEATKQYALGRKKNGEYFGSVAEIIRNHQHSPLVLIDSON 441
 RESULT 7
 AAAS9866
 ID AAAS9866 standard; Protein; 449 AA.
 AC
 XX AAAS9866;
 XX
 DT 20-NOV-1998 (first entry)
 XX
 DE Amino acid sequence of the human BLNK-2 protein.
 XX
 KW Human; BLNK-2; B cells linker protein-2; apoptosis; Grb2; PLC-gamma;

KW SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; Sos;
 KW ras pathway; GDP; GTP; calcium pathway; antigen.
 XX
 OS Homo sapiens.
 XX
 PN WO9832852-A1.
 XX
 PD 30-JUL-1998.
 XX
 PF 23-JAN-1998; 98WO-US01394.
 XX
 PR 17-MAR-1997; 97US-0819013.
 XX
 PR 24-JAN-1997; 97US-0788322.
 XX
 PA (UNITW) UNIV WASHINGTON.
 XX
 PI Chan AC, Fu C;
 XX
 DR WPI: 1998-427948/36.
 XX
 DR N-PSDB; AAY41902.
 XX
 PT Human B cell linker proteins - useful in the treatment of diseases
 XX involving increased or decreased apoptosis
 PS
 XX Disclosure; Fig 3; 56pp; English.
 CC This is the amino acid sequence of the human BLNK-2 (B cells linker
 CC protein-2) protein, used the treatment of diseases involving the
 CC increase and decrease of apoptosis, in the method of the invention.
 CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
 CC intracellular calcium levels and Grb2 is an adapter molecule
 CC containing two SH3 domains that mediate its interaction with the
 CC guanine nucleotide exchange factor. Son of Sevenless (Sos) which in
 CC turn activates the ras pathway by facilitating the exchange of GDP for
 CC GTP on the ras molecule. Activation of both ras and calcium pathways
 CC are required for efficient B cell antigen receptor function. BLNK
 CC binding proteins can be used to identify BLNK proteins in a target
 CC sample.
 XX
 XX Sequence 449 AA:
 SQ
 Query Match 13.8%; Score 319; DB 19; Length 449;
 Best Local Similarity 25.0%; Pred. No. 2.7e-19;
 Matches 99; Conservative 58; Mismatches 139; Indels 100; Gaps 12;
 QY 64 NSNDVDEPDEPQILKAMPKILPAPRIQSEVADTYFDQMEAP---LLLPKAVST 120
 DB 84 NADDSYEPPEVE---QETRPVHPALPFAFGELYDNRSSO--RHSPFSKTLSPKSPWPS 137
 QY 121 EROTRDVRMTQLEEVDPKTFKDVRSQRFKGYTKINKTLPPEPAITLPPKYYOPLPPA 180
 DB 138 EKARLTSTLPALATALOKRQY-----PPKPRGLLEADYVVPV 175
 QY 181 PPESSAYFAKPTTFPEVQKPRORSKADFSVLGAEEESHQTKPSSCPSSQNT--- 237
 DB 176 EDNDENYIHPTESSPPPEKAPM-----VNRSTKPMNSPTPASPPTGASG 219
 QY 238 -----QKSPPAIASSYMP--GKH-----SIQARHTQSM--QKCPAORCOAA- 276
 DB 220 RNSGAMETKSPPPAAPSPLPRACKKPTTPLKTPPVASQANASVCGEKPIPAERHRGSS 278
 QY 277 -----ASHPRMLPYENTNSEKRPD-----TKPDEKVMQNE 308
 DB 279 HROEAQSPVFPFPAQKQIHKQPIPLPRTBEGNPTVDGPLPSFSSNSTISQDAVLCCKP 338
 QY 309 WYIEYSQAVEDVLMKENKDGTLVRCSTKSKAEPYLVVFGNKYVANKIRFLESNQ 368
 DB 339 WYAGACDRKSAEALHRSNKDGSFLIRKSSGHDSCQPYTLVVFENKRYNIPVRIEATK 398
 QY 369 QFALGTGLRGNEMFDSVEDIIEHTYTPILLIDKD 404
 DB 399 QYALGRKKNGEYFGSVAEIIRNHQHSPLVLIDSON 434

```

RESULT 8
AAW59865
ID AAW59865 standard; Protein; 456 AA.
XX
AC AAW59865;
XX
DT 20-NOV-1998 (first entry)
XX
DE Amino acid sequence of the human BLNK-1 protein.
XX
KW Human; BLNK-1; B cells linker protein-1; apoptosis; Grb2; PLC-gamma;
KM SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; SOS;
KW ras pathway; GDP; GTP; calcium pathway; antigen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 327
FT note= "encoded by AGC"
XX
PN W09832852-A1.
XX
PD 30-JUL-1998.
XX
PF 23-JAN-1998; 98WO-US01394.
XX
PR 17-MAR-1997; 97US-0819013.
PR 24-JAN-1997; 97US-0788322.
XX
PA (UNITW ) UNITV WASHINGTON.
PI Chan AC, Fu C;
XX
DR MPI: 1998-427948/36.
DR N-PSSD: AAV41901.
XX
PT Human B cell linker proteins - useful in the treatment of diseases
PT Involved increased or decreased apoptosis
XX
Claim 4; Fig 1; 56pp: English.
XX
XX
This is the amino acid sequence of the human BLNK-1 (B cells linker
protein-1) protein, used the treatment of diseases involving the
increase and decrease of apoptosis, in the method of the invention.
BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
intracellular calcium levels and Grb2 is an adapter molecule
containing two SH3 domains that mediate its interaction with the
guanine nucleotide exchange factor, Son of Sevenless (SOS) which in
turn activates the ras pathway by facilitating the exchange of GDP for
GTP on the ras molecule. Activation of both ras and calcium pathways
are required for efficient B cell antigen receptor function. BLNK
binding proteins can be used to identify BLNK proteins in a target
sample.
XX
SQ Sequence 456 AA;
XX
Query Match 13.8%; Score 319; DB 19; Length 456;
Best Local Similarity 25.0%; Pred. No. 2.7e-19;
Matches 99; Conservative 58; Mismatches 139; Indels 100; Gaps 12.
OY 64 NSNNYDEDEFOLKAMPKMLTPARPQIESPYADTRYQDMMEAP---LLPPRASVST 120
DB 91 NADDSIEPPVE-----QETRPYNPALPFRAGETIDNRSSQ--RHSPPSKTLPSPWS 144
OY 121 EQGTDRVMTOLEEVDKPTFKVDRSORFGKYTKINKTRPLDPAPPALILPKKYOLPLPA 180
DB 145 EKARTLTSLPALTAIAOKROY-----PPRKGLDEDAIYVPV 182
OY 181 PPEESATFARKPTPEYQGRPROASAKDFSNVLCAEBSHHQTPRESSCSNONT--- 237
DB 183 EDNDENYHNPTESSSPPEKAM-----VNSTPNSSTPASPGTIASG 226

```

```

0Y 238 -----OKSPATIASSTYMP--GKH-----SIGARHDTGSM-----QHCPRMORCOAA- 276
      ||||| : : : : : || : : : : : || : : : : :
Db 227 RNSGMETKSPPP--AAPSPLPRAGKPTPTPLKTTTPVASQOAMASSVCEEKPIPERHRHGSS 285
      ||||| : : : : : || : : : : : || : : : : :
0Y 277 ----ASHSPRLPYENTSEKPPD-----TKPDEKDYQNE 308
      || : : : : : || : : : : : || : : : : :
Db 286 HROEAVOSVPRPPAQKQIHOKPIPLPRTPTEGNGPTVDOGPLPIPSNSTISEDEGVLYCKP 345
      ||||| : : : : : || : : : : : || : : : : :
0Y 309 WTIGEYSQOAVEDVLMKENKDGTEFLVNRDCSTKSAEPPYLVPEYGNKVYVWKTREFLESNO 368
      || : : : : : || : : : : : || : : : : : || : : : : :
Db 346 WYAGACDCKSKAEEMALHRSNKGSGSEFLIRKSSGHDSCOPTYTLVFNKRVYNIIPREIATK 405
      ||||| : : : : : || : : : : : || : : : : : || : : : : :
0Y 369 OFALGTGLRGNMEDPSVDILIEHTTYPPIILLIDKQD 404
      ||||| : : : : : || : : : : : || : : : : : || : : : : :
Db 406 QYALGRKNGEYFQSVAILIRNHQSHPLVILDSQN 441
      ||||| : : : : : || : : : : : || : : : : : || : : : : :

```

RESULT 9			
AAW59867			
ID	AAW59867 standard; Protein; 457 AA.		
XX			
XX			
AC	AAW59867;		
XX			
DT	20-NOV-1998 (first entry)		
XX			
DE	Amino acid sequence of the mouse BLNK protein.		
XX			
KW	Mouse; BLNK; B cells linker protein; apoptosis; Grb2; PLC-gamma;		
RW	SH domain; guanine nucleotide exchange factor; Son of Sevenless; SOS;		
KW	ras pathway; GTP; GTP; calcium pathway; antigen.		
XX			
OS	Mus sp.		
PN			
XX	W09832852-A1.		
XX			
PD	30-JUL-1998.		
XX			
PE	23-JAN-1998; 98WO-US01394.		
XX			
PR	17-MAR-1997; 97US-0819013.		
PR	24-JAN-1997; 97US-0788322.		
XX			
PA	(UNIW) UNIV WASHINGTON.		
PI			
XX	Chan AC, Fu C;		
XX			
DR	WPI; 1998-427948/36.		
DR	N-PSDB; AAW41903.		
XX			
PT	Human B cell linker proteins - useful in the treatment of diseases		
PT	Involving increased or decreased apoptosis		
XX			
PS	Disclosure; Fig 5; 56pp; English.		
XX			
CC	This is the amino acid sequence of the murine BLNK (B cells linker		
CC	protein) protein, used the treatment of diseases involving the		
CC	increase and decrease of apoptosis, in the method of the invention.		
CC	BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates		
CC	intracellular calcium levels and Grb2 is an adapter molecule		
CC	containing two SH3 domains that mediate its interaction with the		
CC	guanine nucleotide exchange factor, Son of Sevenless (SOS) which in		
CC	turn activates the ras pathway by facilitating the exchange of GTP for		
CC	GTP on the ras molecule. Activation of both ras and calcium pathways		
CC	are required for efficient B cell antigen receptor function. BLNK		
CC	binding proteins can be used to identify BLNK proteins in a target		
CC	sample.		
XX			
XX			
SQ	Sequence 457 AA;		
Query Match	12.4%; Score 287.5; DB 19; Length 457;		
Best Local Similarity	25.5%; Pred. No. 1.6e-16;		
Matches 106; Conservative	53; Mismatches 139; Indels 117; Gaps		
16;			

QY 66 NNDYDDEPQOLKAMPMSKILPAPLOESEYADTRYFODMEAPLILPPKASVSTEROTR 125
 Db 69 DSDYENPDEH-----SDSEMYVMPAEETGDSYE-----PPPA-----EQQTR 106
 QY 126 DVR-----MTQLEEVKPKPEKDVRSORFK-GFKYTKINKT--PLPPRPAITLP-----KK 173
 Db 107 VVHPALPPTRGKYVNRK-----SQRHSPFFSKTLPSPKSPMSAKARLASTLPAPNSLOK 161
 QY 174 YQPLPAPP-----ESSSAVFAPKPTPEVOGRGRORS-----AKDFSRYLGAEEESHOT 224
 Db 162 PQ-VPPKRDLLDEADYVPEVDENYIHRRESSPPAEAKAPMNRSTKRNSSKHS 220
 QY 225 KPES-----SCPSSNQNTQKSP-----PALASSYMPGKHSIQAR 259
 Db 221 PPGTVAGRNSGVWDSKSLPAPSPILPRAGKKRPAITPLKTPPPLPNAVNCEKRPVPAE 280
 QY 260 DHTGSMQHPQRCQAASH-----SPRMLPYENTNSKRPPT-----297
 Db 261 RHRGS-----SHRODTVQSPVFPPTQKRVHQKPVLPFRFPAGSPADGPPH 327
 QY 298 -----KPEKDYQWQNEWYIGEYSRQAVEDVLMKENKDGTFLYVDCSTKSKAEPPYLV 349
 Db 328 SPFFNLTFADQEGELLGKRWYAGACDRKAEEALHRSNKDGSFLIKSKFGHDSKQPTTLV 387
 QY 350 VFYGNKYVNVKIRPLESNOQFALGTGLRGNEFMFDSVEDIIEHYTYFPILLIDGKD 404
 Db 368 AFENKRVYNIPIVRFIEATQYALGKKKNGEYGSVLEIVNSHONPLVLIDSON 442

RESULT 10

AAW70587
 ID AAW70587 standard; Protein; 95 AA.

AAW70587;

28-JAN-1999 (first entry)

Human SLP-76 SH2 domain.

Human: SLP-130; SLP-76; SH2 domain; leukocyte protein; allergy; cancer;
 phosphoprotein; autoimmune disease; T-cell antigen receptor; TCR;
 immunogen; human immunodeficiency virus; HIV; infectious disease;
 viral; bacterial; fungal; parasitic infection; diabetes mellitus;
 arthritis; dermatitis; asthma; ulcerative colitis; Crohn's disease;
 transplant rejection; graft vs. host disease.

Homo sapiens.

CA2222823-A.

23-JUN-1998.

22-DEC-1997; 97CA-2222823.

03-APR-1997; 97US-0832222.

23-DEC-1996; 96US-0774061.

(IOWA) UNIV IOWA RES FOUNDED.

Hendricks-Taylor RL, Koretzky GA, Motto DG, Musci MA;

WPI, 1999-000127/01.

Nucleic acid encoding SLP-130 protein - useful e.g. in treatment of
 allergy, auto-immune disease and cancer

Example 1: Page 54; 67pp; English.

The present sequence represents the human SH2 domain of the SH2
 containing leukocyte protein of 76 kd (SLP-76). SLP-76 mRNA is expressed
 exclusively in peripheral blood leukocytes, spleen and thymus and its
 over expression augments T-cell antigen receptor (TCR) mediated signals
 that lead to induction of IL-2 gene promoter activity. SLP-130 (SLP-76

associated protein) interacts with SLP-76, is expressed in haematopoietic
 cells and is a substrate for TCR induced activation of a promoter
 containing 3 NFAT sites in a T-cell line and blocks the augmentation of
 this promoter. Isolated SLP-30 may be used as an immunogen to generate
 antibodies that bind SLP-130. Modulation of SLP-130 activity may be
 used to treat immunodeficiency (e.g. human immunodeficiency virus (HIV)
 infection), infectious diseases (e.g. viral, bacterial, fungal and
 parasitic infections), cancer, autoimmune diseases (e.g. diabetes
 mellitus, arthritis, dermatitis, asthma, ulcerative colitis and Crohn's
 disease), transplant rejection, graft vs. host disease and allergies.
 Modulation that results in increased T-cell activation can be used in
 the in vitro production of T-cell cytokines and modulation may be used in
 vaccinations to promote T-cell response to an antigen.

Sequence 95 AA:

Query Match 11.2%; Score 260; DB 20; Length 95;
 Best Local Similarity 53.2%; Pred. No. 4.6e-15;
 Matches 50; Conservative 16; Mismatches 28; Indels 0; Gaps 0;

QY 308 EWYIGEYSRQAVEDVLMKENKDGTFLYVDCSTKSKAEPPYLVYFGNKYVNVKIRPLESN 367
 Db 2 EWYVSTTRPEAFALRKINQDSTFLYVDCSTKSKTTTNPVLMVLYDKYINIRIQKES 61

QY 368 QQFALGTGLRGNEFMFDSVEDIIEHYTYFPILLID 401
 Db 62 QVYLLGTGLRGKEDFLVSQDIIIDYFRKMPILLID 95

RESULT 11

AAG03152
 ID AAG03152 standard; Protein; 84 AA.

AAG03152;

06-OCT-2000 (first entry)

Human secreted protein, SEQ ID NO: 7233.

Human: 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 gene therapy; chromosome mapping.

Homo sapiens.

EP1033401-A2.

06-SEP-2000.

21-FEB-2000; 2000EP-0200610.

26-FEB-1999; 99US-0122487.

(GEST) GENSET.

Dumas Milne Edwards J, Duclert A, Giordano J;

WPI: 2000-500381/45.

N-PSDB; AAC03158.

New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 obtaining cDNAs and genomic DNAs that correspond to 5' ESTs and for
 diagnostic, forensic, gene therapy and chromosome mapping procedures -
 Claim 13; SEQ ID 7233; 71pp + CD-ROM; English.

The present sequence is a polypeptide encoded by one of a large number
 of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 were prepared from total human RNAs or poly(A)⁺ RNAs derived from 30
 different tissues. EST sequences usually correspond mainly to the 3'
 untranslated region (UTR) of the mRNA because they are often obtained
 from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 those cases where longer cDNA sequences have been obtained, the full 5'

CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5' ends and can therefore be used to obtain full length cDNAs and genomic DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and chromosome mapping procedures. They are used to obtain upstream regulatory sequences and to design expression and secretion vectors.

XX Sequence 84 AA;

Query Match Best Local Similarity 9.2%; Score 213; DB 21; Length 84;

Matches 41; Conservative 17; Mismatches 26; Indels 0; Gaps 0;

QY 1 MTSQGNKRTKRGFDLNFQVSLKNSRPSLSAKGRCRAVLEPLDHRNLGAVPGG 60

DB 1 MNRQGNKRTKRGSDNLKFNQSLPKNSRPSRINSATGQYGRMKPLLDRENFVAALDG 60

QY 61 EKCSNNDYEPERFQLKAPSMK 84

DB 61 AKGSHDXXDDPELRMEETWQSIK 84

RESULT 12

ABB38147 ID ABB38147 standard; Peptide: 46 AA.

AC ABB38147;

DT 04-FEB-2002 (first entry)

DE Peptide #5653 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN W0200157277-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001MO-US00669.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483447/52.

DR Human genome-derived single exon nucleic acid probes useful for

PT analyzing gene expression in human fetal liver.

XX Claim 27; SEQ ID NO 30782; 639pp + sequence listing; English.

CC The invention relates to a single exon nucleic acid probe for

CC measuring human gene expression in a sample derived from human foetal

CC liver. The single exon nucleic acid probes may be used for predicting,

CC measuring and displaying gene expression in samples derived from human

CC fetal liver. The present sequence is a peptide encoded by a single exon

CC nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 46 AA;

XX Query Match 9.2%; Score 212; DB 22; Length 46;

Best Local Similarity 89.1%; Pred. No. 2.7e-11; Matches 41; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 329 DGTFLVRDCSTKSKAEPYLVAFYGNKYVYKIRPLESNOQFALGT 374

DB 1 DGSFLVRDCSTKSKAEPYLVAFYENKYVYKIRPLERNOQFALGT 46

RESULT 13

ABB23354 ID ABB23354 standard; Protein: 46 AA.

XX ABB23354;

DT 23-JAN-2002 (first entry)

DE Protein #5353 encoded by probe for measuring heart cell gene expression.

XX Human; gene expression; heart; microarray; vascular system;

KW cardiovascular disease; hypertension; cardiac arrhythmia;

XX congenital heart disease.

OS Homo sapiens.

PN W0200157274-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001MO-US00666.

PR 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488899/53.

DR Single exon nucleic acid probes for analyzing gene expression in human

PT hearts.

XX Claim 15; SEQ ID NO 25124; 530pp; English.

CC The present invention relates to single exon nucleic acid probes for

CC measuring human gene expression in a sample derived from human heart (see

CC ABA21535-ABA41305). The present sequence is a protein encoded by one such

CC probe. The probes may be used for predicting, measuring and displaying

CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,

CC diagnosing, grading, staging, monitoring and prognosing diseases of the

CC human heart and vascular system e.g. cardiovascular disease.

CC hypertension, cardiac arrhythmias and congenital heart disease.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 46 AA;

XX Query Match 9.2%; Score 212; DB 22; Length 46;

XX Best Local Similarity 89.1%; Pred. No. 2.7e-11;

XX Matches 41; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 329 DGTFLVRDCSTKSKAEPYLVAFYGNKYVYKIRPLESNOQFALGT 374

DB 1 DGSFLVRDCSTKSKAEPYLVAFYENKYVYKIRPLERNOQFALGT 46

```

RESULT 14
AA058783
ID AA058783 standard: Protein; 46 AA.
XX
AC AA058783;
XX
DE 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30888.
XX
KW Human; brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 30888; 650bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 46 AA;
XX
Query Match 9.2%; Score 212; DB 22; Length 46;
Best Local Similarity 89.1%; Pred. No. 2.7e-11;
Matches 41; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 329 DGTFLVDCSTKSKAEPLYLVFYGNKYVNVKIRFLESNOOFALGT 374
||:||||||||| ||||| ||| ||||||||||| |||||||
Db 1 DGTFLVDCSTKSKAEPLYLVFYGNKYVNVKIRFLESNOOFALGT 46

```

```

OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 31602; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukaemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 46 AA;
XX
Query Match 9.2%; Score 212; DB 22; Length 46;
Best Local Similarity 89.1%; Pred. No. 2.7e-11;
Matches 41; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 329 DGTFLVDCSTKSKAEPLYLVFYGNKYVNVKIRFLESNOOFALGT 374
||:||||||||| ||||| ||| ||||||||||| |||||||
Db 1 DGTFLVDCSTKSKAEPLYLVFYGNKYVNVKIRFLESNOOFALGT 46

```

Search completed: April 21, 2003, 12:36:24
Job time : 37.4007 secs

Db 91 NADSEYEPPEV---OETRPVHPALPFARGEYIDNRSSQ--RHSPPEFKTLPSKPSWPS 144
QY 121 EROTRDVMTOLEEVYDKPTFKVRSQRKGFYTKINTPLPPRPALTLPRKYYPLRPA 180
Db 145 EKARITSLPALALOKQY-----PPKPKGLLEADYVVPV 182
QY 181 PPESSAYFAPKPTPEVQGRFORSAKDFSVLGAEEESHQTPRESSCSNONOT--- 237
Db 183 EDNDENYIHPTESSSPPEKAPM-----VNRSTKPNSTASPPGTASG 226
QY 238 -----OKSPPALSSSYMP--GKH-----SIQARDHTGSK---QHCPAQCOQA- 276
Db 227 RNSGAMETKSPPP--AAPPLPRACKPTPLKTPVVASQONASSVCEEKPIPAERHRGSS 285
QY 277 ---ASHSPRLPYENTNSEKDP-----TKRDEKDYQNE 308
Db 286 HROEVQSPVPPPAQKQHQKPIPLPRTGEGNPTVDGDLPIFSSNSTISSQEGVLCKP 345
QY 309 WYIEYSHQAVEVLMKENDGTFLVRDCTSKKAEPYLVVYFYNKYKIRPLESNO 368
Db 346 WYAGACDRKSAEALHRSNKDGSFLIRKSSGHSKQPTLVVFFMKRYNIPVRIEATK 405
QY 369 QFALGTGLRGNEMDSVEDIIEHTYFILLIDGKD 404
Db 406 QYALGRKKNGEYFGSAEIIIRNHQHSPLVLIDSQN 441

RESULT 2

US-08-729-416C-1

Sequence 1, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREO
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 474 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-1

Query Match 7.1%; Score 165.5; DB 3; Length 474;
Best Local Similarity 23.2%; Pred. No. 5.5e-07;
Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;
QY 13 GFDDLRFQNVSLKNSRWPSSLSKAGRCRAVLEPLDRHRNLGAGVPGEGKCNNDY--- 69

Db 94 GKSNIQFAGMSLITSTASLNLRTPDSSKQIT--ANHHMRISIFASGSDP--DTTDYVAX 149
QY 70 --EDP---EFOLLKAMPMSKILTPARPIOESEYADTRYFODMEAPLLIPKASVSTERO 123
Db 150 VAKDVPNRRACHILLECCG---LAQDVIGSIGQAFELFKQYLQCPPIIP--ALHDMQ 203
QY 124 TRDVMTOLEEVYDKPTFKVRSQRKGFYTKINTPLPPRPALTLPRKYYPLRPAPE 183
Db 204 SLDEPMT-EGSD-----GSDHPYNSIPSKMPPGGFLDTRKPRHAP-- 247
QY 184 ESSAYFAPK-----PTPE-VORGPFORSAKDFSVLGAEEESHQTPRESSCP 231
Db 248 -DTAQFACKEQTYVGRHLGDTFGEDMQOTPLRQSSD-----ISTPEGKLHVAFTGAP 302
QY 232 SSNONTOKSPPALASSSYMPGKHSIQARDHTGSMOHCAPQCOQAASHSPR-----MLPY 286
Db 303 -TYVNTQOIPP-----QAWPA--AVSSAESPPKKDLFDMPKF 336
QY 287 ENTNSEK-----DPTKPDEK---DYQNEWYIGEYSROAVEDVLMKE 326
Db 337 EDALKNQPLGVLKSAASVECI SPVSPRAPDAKMLELQAEETVYQGENSRKEAEGL- 393
QY 327 NKDGTFLVRDCTSKKAEP--YVLVYFYNKYKIRPLESNOQFALGTGLRGNEMFDS 384
Db 394 EKDDDELVR-----KSTTPGSEFVLGMHNGQAKHLLVDPE-----GITRTDRVFD 442
QY 385 VEDIIEHY--TYFILLIDGKDAARRKCYLTQPL 418
Db 443 ISHLINHLLESSLPV-----SAGSELC-LQOPV 470

RESULT 3

US-08-729-416C-7

Sequence 7, Application US/08729416C
Patent No. 6013767
GENERAL INFORMATION:
APPLICANT: NAKAMURA, TAKESHI
TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
TITLE OF INVENTION: THEREOF, AND ANTIBODY THEREO
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3918
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/729,416C
FILING DATE: 11-OCT-1996
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PERRY, GLENN J.
REGISTRATION NUMBER: 28458
REFERENCE/DOCKET NUMBER: 7898/225948
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-861-3000
TELEFAX: 202-822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 594 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-729-416C-7

Query Match 7.1%; Score 165.5; DB 3; Length 594;
 Best Local Similarity 23.2%; Pred No. 7.6e-07;
 Matches 106; Conservative 60; Mismatches 161; Indels 129; Gaps 25;

13 GGGDLRFQVNSLLKNRSPSLSSAKGRCAVLEPLDHRNRLAGVGGKCSNNNDY--- 69
 214 GKSNIQFAGMSISLITSTASLNRTPDSKQII--ANHHKRSISFASGSDP--DTTDYVAY 269
 70 --EDP-----EFOILKAMPMSKILIPARIOSEVADTRRYQDMMEAFLILPKRASVSTENQ 123
 270 VAKDPVNRRAACHILECCDG--LAQDVISIGAFELRKYLOCPKIP--ALHDMQ 323
 124 TRDVRMTQLEEVDPKPFKDVRSQRFKFKYTKINKTPPLPPRPAITLPKKYOLPPAPPE 183
 324 SLDEPTE--EEDG-----GSDHPYNSIPSKMPRPGGLDRLRKRPRIAP-- 367
 184 ESSAYAPK-----PTFPE-VQGRPROSAKDFSRVLGAEEESHQTPRESSCP 231
 368 -DTAQFAGKEQTYGGRHLGDTFGEEDMOQTPLRQGSDD--IYSTPEGLHVAPTGEAP 422
 232 SSNQNTQKSPPAIASSYMPGKHSIQARHDTGSMOHCAPQCOAASHSPR-----MLPY 286
 423 -TYVNTQQLP-----QAMPA-AVSSAESSPRKLFDMPKP 456
 287 ENTNSEKP-----DPTKDEK---DVMQNEWYIGESQAVEDYLMKE 326
 457 EDLKNQPLGPRVLSKAASVCEISPSRAPADAKMLELAETWYQGMSEKAEGL-- 513
 327 NKGGTLVRDCSTKSKAEP--YVLVVFYGNKYVNVKIRPLESMQOALGTLGNEMFDS 364
 514 EKGGDELVR---KSTNTPESFVLGTGMHNGQAKHLLLVDE-----GTRTRKDRVFD 562
 385 VEDIIEHY--TYPIILLIDGDKAARRKOCYLOPL 418
 563 ISHLINHLESLPIV-----SAGSELC-LQQPV 590

RESULT 4

US-08-807-342B-2
 ; Sequence 2, Application US/08807342B
 ; Patent No. 6077686
 ; GENERAL INFORMATION:
 ; APPLICANT: Der, Channing
 ; APPLICANT: O'Bryan, John P.
 ; APPLICANT: Pawson, Anthony
 ; TITLE OF INVENTION: No. 6077686el SHC Proteins
 ; NUMBER OF SEQUENCES: 18
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Howson and Howson
 ; STREET: Spring House Corporate Cntr., P.O. Box 457
 ; CITY: Spring House
 ; STATE: Pennsylvania
 ; COUNTRY: USA
 ; ZIP: 19477
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/807,342B
 ; FILING DATE: 28-FEB-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/012,516
 ; FILING DATE: 29-FEB-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Bak, Mary E.
 ; REGISTRATION NUMBER: 31,215
 ; REFERENCE/DOCKET NUMBER: MTS2USA
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 215-540-9200
 ; TELEFAX: 215-540-5818

INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 474 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-807-342B-2

Query Match 6.1%; Score 141.5; DB 3; Length 474;
 Best Local Similarity 23.0%; Pred. No. 8.7e-05;
 Matches 101; Conservative 66; Mismatches 177; Indels 95; Gaps 24;

13 GGGDLRFQVNSLLKNRSPSLSSAKGRCAVLEPLDHRNRLAGVGGKCSNNNDY--- 69
 94 GKSNIQFAGMSISLITSTASLNRTPDSKQII--ANHHKRSISFASGSDP--DTTDYVAY 149
 70 --EDP-----EFOILKAMPMSKILIPARIOSEVADTRRYQDMMEAFLILPKRASVSTENQ 123
 150 VAKDPVNRRAACHILECCDG--LAQDVISIGAFELRKYLOCPKIP--ALHDMQ 203
 124 TRDVRMTQLEEVDPKPFKDVRSQRFKFKYTKINKTPPLPPRPAITLPKKYOLPPAPPE 183
 204 SLDEPTE--EEDG-----DHPYNSVPTKMPRPGGLDRLRKRPRIAP-- 247
 184 ESSAYAPK-----PTFPE-VQGRPROSAKDFSRVLGAEEESHQTPRESSCP 231
 248 -EAAQFAGKEQTYGGRHLGDTFGEEDMOQRAPTRQSLDIYST--ABGKTH--WVPGERP 302
 232 SSNQNTQKSP-----AIASSYMPGKHSIQARHDTGSMOHCAPQCOAASHSPMLPY 286
 303 -TYVNTQPPRPQWPAATSTESSPPRKLFDMPKPFEDALNQPLSKAASVCEISPV 361
 287 ENTNSEKPDPKDEK---DVMQNEWYIGESQAVEDYLMKENKDGTLVRDCSTSKA 343
 362 -----TPRADAKMLELAETWYQGMSEKAEGL--REGGDELVR---KSTT 406
 344 EP--YVLVVFYGNKYVNVKIRPLESMQOALGTLGNEMFDSVEDIIEHY--TYPIILL 399
 407 NPESFVLGTGMHNGQAKHLLLVDE-----GTRTRKDRVFDISHLINHLESLPIV 458
 400 IDGDKAARRKOCYLOPL 418
 459 -----SAGSELC-LQQPV 470

RESULT 5

US-08-353-550-6
 ; Sequence 6, Application US/08353550
 ; Patent No. 5744313
 ; GENERAL INFORMATION:
 ; APPLICANT: Williams, Lewis T.
 ; APPLICANT: Cavanaugh, William M.
 ; TITLE OF INVENTION: No. 5744313el Protein Domain Which Binds
 ; TITLE OF INVENTION: Tyrosine Phosphorylated Proteins
 ; NUMBER OF SEQUENCES: 9
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend Kourie and Crew
 ; STREET: One Market Plaza, Stewart Tower, Suite 2000
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94105
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/353,550
 ; FILING DATE: 09-DEC-1994
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Dow, Karen B.

APPLICANT: Ullrich, Axel
APPLICANT: Gishlitzky, Mikhail
APPLICANT: Sures, Irman G.
TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
TITLE OF INVENTION: TYROSINE KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York,
STATE: NY
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: PASTSEQ Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/426,509A
FILING DATE: 21-APR-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/232,545
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-0074-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 675 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: No. 6326469e
US-08-426-509A-4

Query Match 5.7%; Score 133; DB 4; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.00088;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;
QY 166 PATLPKRYOPLPAPPESSAYFAKPTPE-VQGRPR-----QSAKDFRYLGA--- 216
DB 141 PGCLTMEAYANLHTAVNEKHRV-----PTFPDRVLKIPRAVPVLKMDAPSSSTLLAQYDN 196
QY 217 EESHQTKPESSCPSSNQNTOKSPAIASS-----SYMPG-----KHSIOAR 259
DB 197 ESKKNYGSQPPSSSTSLAQYDSNSKKIYGSOPNFMQYIIPREDPPDMQVRAKLKSSSSSE 256
QY 260 DHGSMOHCAPQOQAASHSPRLPYENTNSEKPDPTKPEKDVQWQENYIGEYSQAV 319
DB 257 DVASSNQ-----ERNVNTTSKISWEFPSSSE-----EENLDYDWFAGNISRSQS 306
QY 320 EDVLMKENKDGFLVRDCSTKSKAEPYLVLF---YGNKYVNVKIRLESNOQFALGTGL 376
DB 307 EQLLRQKKGEGAFNVRN---SSQVGMVTVSLFSKAVNDKKTGVKHHVHTNAENKLL--YL 361
QY 377 RGNEMFDSVEDIIEHYTY 394
DB 362 AENYCFDSIPKLTHYQH 379

RESULT 10
PCT-US95-05008-4
Sequence 4, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Calveston Drive

APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissensschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
APPLICANT: Germany
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-APR-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-APR-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 675 amino acids.
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-05008-4

Query Match 5.7%; Score 133; DB 5; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.00088;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;
QY 166 PATLPKRYOPLPAPPESSAYFAKPTPE-VQGRPR-----QSAKDFRYLGA--- 216
DB 141 PGCLTMEAYANLHTAVNEKHRV-----PTFPDRVLKIPRAVPVLKMDAPSSSTLLAQYDN 196
QY 217 EESHQTKPESSCPSSNQNTOKSPAIASS-----SYMPG-----KHSIOAR 259
DB 197 ESKKNYGSQPPSSSTSLAQYDSNSKKIYGSOPNFMQYIIPREDPPDMQVRAKLKSSSSSE 256
QY 260 DHGSMOHCAPQOQAASHSPRLPYENTNSEKPDPTKPEKDVQWQENYIGEYSQAV 319
DB 257 DVASSNQ-----ERNVNTTSKISWEFPSSSE-----EENLDYDWFAGNISRSQS 306
QY 320 EDVLMKENKDGFLVRDCSTKSKAEPYLVLF---YGNKYVNVKIRLESNOQFALGTGL 376
DB 307 EQLLRQKKGEGAFNVRN---SSQVGMVTVSLFSKAVNDKKTGVKHHVHTNAENKLL--YL 361
QY 377 RGNEMFDSVEDIIEHYTY 394
DB 362 AENYCFDSIPKLTHYQH 379

RESULT 11
US-08-807-342B-4

Db 232 -ALALTDGSPSLRDLACSLPMDVSGTAPPGD-----YVQADARGP--- 296
Qy 206 SAKDFSRVGAEEESHNOTKPESSCPSSNONOTOKSPPAIASSSYMPCGKHSIOARDHTGSM 265
Db 297 --PDHEHLVYVNTQGLDAPEPEDS-----PKKDLDFMREEDAL 333
Qy 266 QHCPAORCOAAASHSPRLPYENTNSEKPD--PTKPEKDVQWQEWYIGEYSRQAVEDV 322
Db 334 K---LHECSVAGVTAALPLLEDQWPSPTTRAPYAPTEEOLRQEPWYHGRMSRRAERM 390
Qy 323 LMKENKDGFLVRCDSST 339
Db 391 L---RADGDFLVDSVT 404

RESULT 13
US-07-906-349A-5

; Sequence 5, Application US/07906349A
; Patent No. 5434064
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnik, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AND
; NUMBER OF INVENTION: IDENTIFYING TARGET PROTEINS
; TITLE OF INVENTION: TARGET PROTEINS
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Broadway and Nelmark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,349A
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/643,237
; FILING DATE: 18-JAN-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-373-3528
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 724 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-906-349A-5

Query Match 5.5%; Score 128; DB 1; Length 724;
Best Local Similarity 20.6%; Pred. No. 0.0028;

Matches 78; Conservative 45; Mismatches 135; Indels 120; Gaps 14;

Qy 81 PSMKILPARPIQSESYADTRFYQDMMEAPLLLPKAS----- 117
Db 85 PTPKRPRPRPLVPAP-GSSKTEADVEQOALTLPLDLAEQFAPPDIAAPLLIKLVEAIEKKG 143
Qy 118 -----VSTEQTRDVNRMTQLEEVNDKPTFK-----DVRSQRFKGFYTKINKTPPLPPR 165
Db 144 LECSTLYRTQSSSLNLAELRQLDCTPSVDLEMIDVHVLADAFKRYL-----LDLPN 195
Qy 166 PAITLPKRYQPLPPAPRESSAVFAP-----KTFPEVQCGPRQASAKDFSRVGAEE 218
Db 136 PVIIPAIVSEMISTLAPEVQSSSEETIQLKLIRSPSIHQIWLTLQYLKHFKL----- 250

Qy 219 ESHQTKPESSCPSSNONOTOKSPPAIASSSYMPCGKHSIOARDHTGSM----- 265
Db 251 -----SQTSKLNLAARVLSEIFSPMLFFRFSAASSDNENLKIIEILLISTEMN 299
Qy 266 --QHCPAORCOAAASHSPRLPYENTNSEKPDPTFK-----DEQDVQWQEWYIGEY 314
Db 300 EROPAPA-----LP-----PKPKPTTVANNGNMNNLSQNAEYWGDI 338
Qy 315 SROAVEDVLMKENKDGFLVRCDSSTKSAEPYVLVVFQGNKYVVKIRFLESNOOFALGT 374
Db 339 SREEVNEKL-RDTADGTLVADASTKMGD--YTLFLRKGDN--KKLIKIFHRDCKYGRSD 394
Qy 375 GLRGNEMFDSVEDIIEHY 392
Db 395 PL---TFSSVVELINHY 408

RESULT 14
US-08-167-035-2

; Sequence 2, Application US/08167035
; Patent No. 5618691
; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnik, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
; IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: 10036-2711
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,035
; FILING DATE: 16-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Corvuzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 724 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-167-035-2

Query Match 5.5%; Score 128; DB 1; Length 724;
Best Local Similarity 20.6%; Pred. No. 0.0028;

Matches 78; Conservative 45; Mismatches 135; Indels 120; Gaps 14;

Qy 81 PSMKILPARPIQSESYADTRFYQDMMEAPLLLPKAS----- 117
Db 85 PTPKRPRPRPLVPAP-GSSKTEADVEQOALTLPLDLAEQFAPPDIAAPLLIKLVEAIEKKG 143
Qy 118 -----VSTEQTRDVNRMTQLEEVNDKPTFK-----DVRSQRFKGFYTKINKTPPLPPR 165
Db 144 LECSTLYRTQSSSLNLAELRQLDCTPSVDLEMIDVHVLADAFKRYL-----LDLPN 195

[illegible]

```

RESULT 2
US-09-966-955A-2
; Sequence 2, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gena S.
; APPLICANT: Kanter, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Cink-related Gene, MIST- (Mast Cell Immunoreceptor
; FILE REFERENCE: 3053-413US1
; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN FULL-LENGTH MIST cDNA CLONE #8, TRANSLATED
; US-09-966-955A-2

Query Match          56.8%; Score 1316.5; DB 9; Length 443;
Best Local Similarity 62.1%; Pred. No. 5,76-85;
Matches 266; Conservative 46; Mismatches 111; Indels 5; Gaps 4;

QY      1  MTSQGNKRTYEGGDLRFQVNSLKNRSPSSLSAKRCRAVLELPDPHRRNLGVPDG 60
      |  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      16  MNRGNKRRTKTEGGNDLKFQNFSLPKNRSWPRINSATQYOYRMNKLPDMEWENFAAVLDG 75
      |  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      61  EKCSNNDDYDEPFOLKAMPMSKILPARPQSEFVATRYFOODMEAPLLPKASYST 120
      |  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      76  AKGSDDDYDDEPLRMEETWOSTITLPAAPRIKESYADTHYFKAMDDPLPDTDTSTISI 135
      |  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      121  ERQTRVAMTQLEEVDRPTFKDVRQSRQKGGKYYKINKPLPPRPATILPKKYQPLPPA 180
      |  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      136  GQPTWMTQ-TRLEVRDAPISKDVRSQNIKGDAVSRKNNKILPLPPRPDLITLKKYQPLPP- 193
      |  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY      181  PPESSAIVPAKPTFPFYQGRORSKADSRVLCAEEESHQKRPRESSCPSSNONTQKS 240
      |  |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

Db	194	EPBSSRPLSQRHTFPEVQMRQISLRDLSVLEAEKVRPNQRKPESTHLENOQTOI	2533
Qy	241	PRAIASSSYMGKHSIOARDHTGCMOCPAROCSAAASHSP--RMLPYENTNSEKRPDTK	2988
Db	254	PLAISSSSFTTSMNSVQNRHGRGMOQCSPORCQPRASCSPHEMLRKYTKTSRRPREKR	3133
Qy	299	PDEKQWQNEWTIGETSHQAVEDYLAKENKDGTLVYNDOSTKSKAEPTYLVFTGKNVYN	3588
Db	314	SDRKVQVQNEWTIGETSHQAVEDYLAKENKDGSTFLVYNDOSTKSKAEPTYLVAFYEKNVYN	3733
Qy	359	VKIRELESNOQFALGTGLRGNEPMDVEDIIEHTYTPRILIDGKDR--AARRQCYLTPR	417
Db	374	VKIRELENGQFALGTGLRGKEKDSVEDIIEHTKKNPRILIDGKDRGTVNRRQCHLTPR	433
Qy	418	LPLARLL 425	
Db	434	LPLTRHLL 441	

```

RESULT 3
US-09-966-955A-6
; Sequence 6, Application US/09966955A
; Patent No. US2002015563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whltney, Gena S.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Clnk-related Gene, MIST (Mast Cell Immunoreceptor)
; FILE REFERENCE: 3053-4113U51
; CURRENT APPLICATION NUMBER: US/09/966,955A
; PRIOR FILING DATE: 2001-09-28
; PRIORITY FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO: 6
; LENGTH: 353
; TYPE: PRF
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN MIST SPICE VARIANT CLONE #12, TRANSLATED
; OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-6

Query Match      47.9%; Score 1109.5; DB 9; Length 353;
Best Local Similarity 64.4%; Pred. No. 1.5e-70;
Matches 226; Conservative 31; Mismatches 89; Indels 5; Gaps 4;

QY   78 KAPSMKILPAPROESRYADTRYQDMMEARLLLPKASVSTENOTRDVRTOLEEVDK 137
    3 ETWOSIKLPPAPRIKESSEYADTHYFKVMDPLPDRTISISIGPTNNTQ-TRELERYDK 61

QY   138 PTFFDVRSQRFGKGYRTINKTPLEPPRPATTLPPKKYOLPRLPAPESSAYFPKPTEPE 197
    62 PISDVRSQNKGDAVAKNKIPLPPRPLLTLPPKYOPLP-EPESSRPLSQHHTPE 120

QY   198 VQGRPRORSANDFSRVYLGAEEESHQTPRESSCPSNNONTOKSPALAISSYMPGKHSTO 257
    121 VQMRPSQSILRDLSEVLAEAKVPINHQRPETHLLENQTOELPLAISSSSTTTSNHSYQ 180

QY   258 ARDHGTSGMQHCPAORCQAASHSP--RMLPYENTNSEKPDPTKPEDKDVMONENYIGEYS 315
    181 NRDRHGQMOPCSPORCOPASCSPHENILPYKYTSWRPFRPSDRSKOVQHNEWYIGETS 240

QY   316 RQAVEEDLMKKNKDTFLVDCSTKSKEPYLVLVFYGNKYVNVKIRPLESNOGFALGTG 375
    241 RQAAVEAFMKENKGSFLYVRDCSTKSKEEPYLVAVFYENKYNVVKIRPLERNQGFALGTG 300

```


QY 376 LRGNEMFDSVEDIEIHTYPIILIDGDK-AAARKOCYTLOPLIARLL 425
Db 301 LRGEKFDSDVEDIEIHTKPNPILIDGKDTGVHRKCOCHLTOPLEPLRHL 351

RESULT 4
US-09-864-761-38652
; Sequence 38652: Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/226,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/224,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38652
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005599.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4
; OTHER INFORMATION: SWISSPROT HIT: Q60787, EVALU 9.00e-10
US-09-864-761-38652

Query Match 9.2%; Score 212; DB 10; Length 46;
Best Local Similarity 89.1%; Pred. No. 1.3e-08;
Matches 41; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 329 DGFELVDCSTKSKAEPPYLWVEYGNKYVNVKIRPLESNOQFALGT 374
Db 1 DGSFELVDCSTKSKAEPPYLWVEYGNKYVNVKIRPLESNOQFALGT 46

RESULT 5
US-10-186-399-3
; Sequence 3, Application US/10186399
; Patent No. US20020173481A1
; GENERAL INFORMATION:
; APPLICANT: Ekman, Niklas
; APPLICANT: Arighi, Elena
; APPLICANT: Vastrik, Imre
; APPLICANT: Tamagnone, Luca
; APPLICANT: Allalo, Karl
; TITLE OF INVENTION: REGULATION OF VASCULAR ENDOTHELIUM USING BMX TYROSINE
; FILE REFERENCE: 28113/31941A
; CURRENT FILING DATE: 2002-07-01
; PRIOR APPLICATION NUMBER: US 08/320,432
; PRIOR FILING DATE: 1994-10-07
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 3
; LENGTH: 675
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-186-399-3

Query Match 5.7%; Score 133; DB 9; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.13;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

QY 166 PATLPKKYQPLPAPPESSAYAPKPTPE-VQRGPR-----QRANKFSRYLGA--- 216
Db 141 PGCTIMEAYANLHTAVNEKHV---PTPPDVLKIPRAVPYIKMDAPSSSTTLAODYN 196
QY 217 EESHQHTKPESSCPSSNONOTQKSPAIASS-----SYMFG-----KHSIOAR 259
Db 197 ESKKNYSQPPSSSTSLAQYDSNKKIYSGPNPNMGIYREDPPDMQYRKLSSSSSSE 256
QY 260 DHTGMOHCPAORCAASHSPRLPYENTNSEKPDPTKDEKDVONENWYIGEYSQAV 319
Db 257 DVASSNCK-----ERNVNHSTSKISWEPSSSSE-----EENLDYDWFAGNISRSOS 306
QY 320 EDVLMKKNKDGFLVRCSTSKAEPPYLWVEYGNKYVNVKIRPLESNOQFALGTGL 376
Db 307 EQLLRQKKEGAEFWRN---SSQVGYTVSLSKAVNDKGTVAHYHVTNAENKL--YL 361
QY 377 RGNEMFDSVEDIEIHTY 394
Db 362 AENYCFDSIRPLIHYHQH 379

RESULT 6
US-09-977-260-4
; Sequence 4, Application US/09977260
; Publication No. US20020192790A1
; GENERAL INFORMATION:
; APPLICANT: ULRICH, AXEL
; APPLICANT: GISHIKY, MIKHAEL
; APPLICANT: SURES, IRMINGARD
; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
; FILE REFERENCE: 038602/1260
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 08/232,545

;; PRIOR FILING DATE: 1994-04-22
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 675
;; TYPE: PRT
;; ORGANISM: Unknown Organism
;; FEATURE:
;; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
US-09-977-260-4

Query Match 5.7%; Score 133; DB 9; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.13;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

QY 166 PATLPKKYQPLPPAPPESSAYFAKPTPE-VQRCPR-----QSAKDSRVLGA--- 216
DB 141 PGCTLMEAYANLHTAVNEKHRV---PTFPDVLKIPRAVPVLKMDAPSSSTTLAQYDN 196
QY 217 EESHSHQTPRESSCPSSNONTKSPPAIASS-----SYMPC-----KHSIQAR 259
DB 197 ESKKNYGSQPPSSSTSLAQYDSNKKIYGSQPNFMQYIPREDPDMQVRLKSSSSSE 256
QY 260 DHTGSMQHCAPQCAAAASHSPRLPYENTNSEKPDPTKDEKDYQWQNEWYIGEYSRQAV 319
DB 257 DVASSNOK-----ERNVHHTTSKISWEPESSSE-----EENLDYDWMFAGNISRSQS 306
QY 320 EDVLKKNKDGTFILVRCSTKSKAEPLYLVF---YGNKYVNVKIRPLESNOQFALGTGL 376
DB 307 EQLLRQKQKGAFAFWRN---SSQVGMVTVSLFSAVNDKKGTVKHYHHTNAENKL--YL 361
QY 377 RGNEMFDSVEDIIEHYTY 394
DB 362 AENYCFDSIPKLIHYHGH 379

RESULT 7
US-09-977-261-4
;; Sequence 4, Application US/09977261
;; Publication No. US20030054527A1
;; GENERAL INFORMATION:
;; APPLICANT: ULBRICH, AXEL
;; APPLICANT: GISHIZKY, MIKHAIL
;; APPLICANT: SURES, IRMINGARD
;; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
;; FILE REFERENCE: 038602/1259
;; CURRENT APPLICATION NUMBER: US/09/977,261
;; PRIOR FILING DATE: 2001-10-16
;; PRIOR APPLICATION NUMBER: 08/232,545
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 675
;; TYPE: PRT
;; ORGANISM: Unknown Organism
;; FEATURE:
;; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
;; OTHER INFORMATION: kinase 2
US-09-977-261-4

Query Match 5.7%; Score 133; DB 9; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.13;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

QY 166 PATLPKKYQPLPPAPPESSAYFAKPTPE-VQRCPR-----QSAKDSRVLGA--- 216
DB 141 PGCTLMEAYANLHTAVNEKHRV---PTFPDVLKIPRAVPVLKMDAPSSSTTLAQYDN 196
QY 217 EESHSHQTPRESSCPSSNONTKSPPAIASS-----SYMPC-----KHSIQAR 259
DB 197 ESKKNYGSQPPSSSTSLAQYDSNKKIYGSQPNFMQYIPREDPDMQVRLKSSSSSE 256

QY 260 DHTGSMQHCAPQCAAAASHSPRLPYENTNSEKPDPTKDEKDYQWQNEWYIGEYSRQAV 319
DB 257 DVASSNOK-----ERNVHHTTSKISWEPESSSE-----EENLDYDWMFAGNISRSQS 306
QY 320 EDVLKKNKDGTFILVRCSTKSKAEPLYLVF---YGNKYVNVKIRPLESNOQFALGTGL 376
DB 307 EQLLRQKQKGAFAFWRN---SSQVGMVTVSLFSAVNDKKGTVKHYHHTNAENKL--YL 361
QY 377 RGNEMFDSVEDIIEHYTY 394
DB 362 AENYCFDSIPKLIHYHGH 379

RESULT 8
US-09-977-269-4
;; Sequence 4, Application US/09977269
;; Patent No. US20020082037A1
;; GENERAL INFORMATION:
;; APPLICANT: ULBRICH, AXEL
;; APPLICANT: GISHIZKY, MIKHAIL
;; APPLICANT: SURES, IRMINGARD
;; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN TYROSINE KINASES
;; FILE REFERENCE: 038602/1260
;; CURRENT APPLICATION NUMBER: US/09/977,269
;; PRIOR FILING DATE: 2001-10-16
;; PRIOR APPLICATION NUMBER: 08/232,545
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: Patentln Ver. 2.1
;; SEQ ID NO 4
;; LENGTH: 675
;; TYPE: PRT
;; ORGANISM: Unknown Organism
;; FEATURE:
;; OTHER INFORMATION: Description of Unknown Organism: Megakaryocyte
;; OTHER INFORMATION: kinase 2
US-09-977-269-4

Query Match 5.7%; Score 133; DB 10; Length 675;
Best Local Similarity 24.0%; Pred. No. 0.13;
Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

QY 166 PATLPKKYQPLPPAPPESSAYFAKPTPE-VQRCPR-----QSAKDSRVLGA--- 216
DB 141 PGCTLMEAYANLHTAVNEKHRV---PTFPDVLKIPRAVPVLKMDAPSSSTTLAQYDN 196
QY 217 EESHSHQTPRESSCPSSNONTKSPPAIASS-----SYMPC-----KHSIQAR 259
DB 197 ESKKNYGSQPPSSSTSLAQYDSNKKIYGSQPNFMQYIPREDPDMQVRLKSSSSSE 256
QY 260 DHTGSMQHCAPQCAAAASHSPRLPYENTNSEKPDPTKDEKDYQWQNEWYIGEYSRQAV 319
DB 257 DVASSNOK-----ERNVHHTTSKISWEPESSSE-----EENLDYDWMFAGNISRSQS 306
QY 320 EDVLKKNKDGTFILVRCSTKSKAEPLYLVF---YGNKYVNVKIRPLESNOQFALGTGL 376
DB 307 EQLLRQKQKGAFAFWRN---SSQVGMVTVSLFSAVNDKKGTVKHYHHTNAENKL--YL 361
QY 377 RGNEMFDSVEDIIEHYTY 394
DB 362 AENYCFDSIPKLIHYHGH 379

RESULT 9
US-10-081-980B-1
;; Sequence 1, Application US/10081980B
;; Publication No. US20030041337A1
;; GENERAL INFORMATION:
;; APPLICANT: Pfizer Inc.
;; APPLICANT: Gibbs, E. Michael
;; APPLICANT: McEnish, John D.
;; TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form

Query Match	5.6%;	Score 130;	DB 9;	Length 724;
Best Local Similarity	22.0%;	Pred. No. 0.23;		
Matches	89;	Conservative	46;	Mismatches 147;
		Indels	122;	Gaps 19

[illegible]

Db 288 OKKGGAGMVRN---SQMGMTVSLFSKAVNDKKGTVKHHVHTNAENKL--YLAENYC 342
QY 382 FDSVEDIEIHYTY 394
Db 343 FDSIKKLHHYH 355

RESULT 12

US-09-962-929-4
; Sequence 4, Application US/09962929
; Patent No. US20020115058A1

GENERAL INFORMATION:

APPLICANT: Pedersen, Finn S.
APPLICANT: Soerensen, Annette B.
APPLICANT: Nielsen, Anne A.
TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated with A
FILE REFERENCE: A-70004/RMS/DCF
CURRENT APPLICATION NUMBER: US/09/962,929
PRIOR FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 09/668,644
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin version 3.1
SEQ ID NO 4
LENGTH: 724
TYPE: PRT
ORGANISM: Homo sapiens
US-09-962-929-4

Query Match 5.5%; Score 128; DB 10; Length 724;
Best Local Similarity 20.6%; Pred. No. 0.32;
Matches 78; Conservative 45; Mismatches 135; Indels 120; Gaps 14;

QY 81 PSMKILPAPPIQSEYADTRFYQDMMEAPLLPPKAS----- 117
Db 85 PTPKPRPPPLPVPAP-GSSKTEADVEQOALTLPLDAEQFAPDAPDAPLLIKLVEAIKKG 143
QY 118 -----VSTERQTRDVMTQLEEVDPKPTFK-----DVRSORFKGFYTKINKTLPPLPR 165
Db 144 LECSTLYRTQSSNLAEIRQLDCTDPSVLEMTDVHLADAFKRYL-----LPLPN 195
QY 166 PATLPKRYQPLPAPPESSATFAP-----KPTPEVQGRPRQSRANDSFVLGAEE 218
Db 196 PVIPAAYSEIISLAPEYQSEEXIQLKILIRSPISIPHQYWLQYLKHFFKL----- 250
QY 219 ESHHQTPRESSCPSSNONTOKSPPAIASSVMPGKHSIQARDHTGSM----- 265
Db 251 -----SQTSSKNLNAARVLSEIFSPMLFRFSASSDNTENLIKVIILISTEWN 299
QY 266 --OHCPAORCOAAASHSPMLPYENTNSEKPDPTKP-----DEKDVQWQNEWYIGY 314
Db 300 ERQAPAP-----LP-----PKPKPTTVANNNGNNMMSIQNMWYIGDI 358
QY 315 SRAOVEVLAKENKDGTFVLRDCSTKSKAEPLYLVVFGKVVNVKTRFLESNOQFALGT 374
Db 339 SREEVNKL-RDTADGFLVRDASTKMGHD-YTLLEKGGN-NKLKIFHRDQKYGFS 394
QY 375 GLRGNEFMFDSVEDIEIHY 392
Db 395 PL-----TFSSVELINHY 408

RESULT 13

US-09-962-929-2
; Sequence 2, Application US/09962929
; Patent No. US20020115058A1

GENERAL INFORMATION:

APPLICANT: Pedersen, Finn S.
APPLICANT: Soerensen, Annette B.
APPLICANT: Nielsen, Anne A.
TITLE OF INVENTION: Methods for diagnosis and treatment of diseases associated with A
TITLE OF INVENTION: Expression of Plk3l1

FILE REFERENCE: A-70004/RMS/DCF
CURRENT APPLICATION NUMBER: US/09/962,929
CURRENT FILING DATE: 2001-09-24
PRIOR APPLICATION NUMBER: US 09/668,644
PRIOR FILING DATE: 2000-09-22
NUMBER OF SEQ ID NOS: 4
SOFTWARE: Patentin version 3.1
SEQ ID NO 2
LENGTH: 724
TYPE: PRT
ORGANISM: Mus musculus
US-09-962-929-2

Query Match 5.5%; Score 127; DB 10; Length 724;
Best Local Similarity 22.0%; Pred. No. 0.38;
Matches 89; Conservative 45; Mismatches 148; Indels 122; Gaps 19;

QY 38 GRCAVLEPLPDHR-RNLGVPGEKCNSSNDYED-PEFOLKAMPKMLPAPPIQES 94
Db 78 GRKR-ISPPTKPPRPPPLPVPAPSSKTEADTEQOALPLDLAEQFAPDAPVAPPLIKL 136
QY 95 EYADTRFYQDMMEAPLLPPKASVSTERQTR-----DVMTQLEEVDPKPTFKDVRSORFK 149
Db 137 EAIKK-----GLECSTLYRTQSS-SNPABLRLDCCOASVDLEMTDVHLADA----FK 187
QY 150 GFYTKINKTPTLP-----PRAITLPKRYQPLPAPPE-ESSATFAPK 192
Db 188 --RYLADLPNEVLPVAVYNEMSLAQELQSPEDCIQLKILIRLPNIPHCWMLQYLLK 245
QY 193 PTPPEVQGRPRQSRAND-----FSRYL-----GAEESSHQTPRESSCPSSNQNT 237
Db 246 HFFK-----LSQASSKNLNAARVLSEIFSPVLPFRPAASSDNTENLIKALIELISTWNE 300
QY 238 QKSPPAIASSSYMFGKHSIQARDHTGSMOHCAPACQAAASHSPMLPYENTNSEKPDPT 297
Db 301 ROPAPAL-----PKKPP 312
QY 298 KP-----DEKDVQWQNEWYIGYVROAVEDVLMKENKDGTFVLRDCSTKSKAEPLY 348
Db 313 KPTTVANNMNNMNSLODAEMWGDISREEVNKL-RDTADGFLVRDASTKMGHD-YTL 370
QY 349 VVEYGNKVNVKIRFLESNOQFALGTGLRGNEFMFDSVEDIEIHY 392
Db 371 TPKKGGN-NKLKIFHRDQKYGFSPL-----TFNSVELINHY 408

RESULT 14

US-09-920-021A-3
; Sequence 3, Application US/09920021A
; Patent No. US20020110800A1

GENERAL INFORMATION:

APPLICANT: KAPLAN, DAVID
APPLICANT: MARSH, H. NICK
TITLE OF INVENTION: USE OF SHP-1 AND SHP-2 TO DETECT
TITLE OF INVENTION: COMPOUNDS INVOLVED IN NEURONAL SURVIVAL
FILE REFERENCE: 08338/016001
CURRENT APPLICATION NUMBER: US/09/920,021A
CURRENT FILING DATE: 2001-08-01
PRIOR APPLICATION NUMBER: US/08/918,157
PRIOR FILING DATE: 1997-08-25
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 593
TYPE: PRT
ORGANISM: Homo sapiens
US-09-920-021A-3

Query Match 5.4%; Score 124.5; DB 10; Length 593;
Best Local Similarity 29.3%; Pred. No. 0.44;
Matches 43; Conservative 21; Mismatches 50; Indels 33; Gaps 5;

QY 295 DPTKPEKDVQWQNEWYIGYVROAVEDVLMKENKDGTFVLRDCSTKSKAEPLYLVVFG- 353

```
Db 106 DPT-----SERWFHGLSGKEAKLLEKCKHGSLVRE--SQSHGDFVLSVRGCD 155
|||
QY 354 -----NKYYNKIRFLESNOQFALGTGLRGNEMEDSVEDITEHTYTFPILIDGKD 404
|||
Db 156 DKGESNDKSKVTHWIRCOE-----LKVDYGGGERFDSLTDLVEHYKKNPMEVETLG-- 207
|||
QY 405 KAARRKQCYLTOPLPLARLLLTQYSSQ 431
|||
Db 208 -----TVLQKQPLNTTRINAAEIESR 229
|||
```

RESULT 15

```
US-10-038-010-22
: Sequence 22, Application US/10038010
: Publication No. US20030040089A1
: GENERAL INFORMATION:
: APPLICANT: HYBRIGENICS
: APPLICANT: Pierre, Legrain
: TITLE OF INVENTION: Protein-protein interactions in adipocyte cells
: FILE REFERENCE: BA767A
: CURRENT APPLICATION NUMBER: US/10/038,010
: PRIOR FILING DATE: 2002-07-23
: PRIOR APPLICATION NUMBER: US 60/259,377
: NUMBER OF SEQ ID NOS: 67
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 22
: LENGTH: 597
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: hSHP2_FL
: LOCATION: (1) (597)
: OTHER INFORMATION:
US-10-038-010-22
```

Query Match

5.4%; Score 124.5; DB 9; Length 597;

Best Local Similarity 29.3%; Pred. No. 0.45; Mismatches 50; Indels 33; Gaps 5;

```
QY 295 DPTKPEDKDVQWQNEWYIGEYSRQAVEDYLMKENKDGTEFLVRDCSTKSKAEPYLVVFG- 353
|||
Db 106 DPT-----SERWFHGLSGKEAKLLEKCKHGSLVRE--SQSHGDFVLSVRGCD 155
|||
QY 354 -----NKYYNKIRFLESNOQFALGTGLRGNEMEDSVEDITEHTYTFPILIDGKD 404
|||
Db 156 DKGESNDKSKVTHWIRCOE-----LKVDYGGGERFDSLTDLVEHYKKNPMEVETLG-- 207
|||
QY 405 KAARRKQCYLTOPLPLARLLLTQYSSQ 431
|||
Db 208 -----TVLQKQPLNTTRINAAEIESR 229
|||
```

Search completed: April 21, 2003, 12:39:41
Job time : 18.5549 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:33:12 ; Search time 16.6276 seconds
(without alignments)
2515.001 Million cell updates/sec

Title: US-09-856-061-2

Perfect score: 2316

Sequence: 1 MTSQGNKRTKEGFGDLRPO.....QPLPLALLLQYSSQALHE 435

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283224 segs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR73:*
2: PIR1:*
3: PIR2:*
4: PIR4:*

Pred. No. is the number of results predicted by chance a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	412.5	17.8	533	2 B56110	tyrosine phosphop
2	398	17.2	533	2 A56110	tyrosine phosphop
3	154.5	6.7	663	1 TVMYRR	protein-tyrosine k
4	146	6.3	728	2 H59435	3-phosphatidylinos
5	142.5	6.2	723	2 B38749	3-phosphatidylinos
6	138.5	6.0	443	2 T27877	hypothetical prote
7	134.5	5.8	480	2 JC7552	Shb-like adapter p
8	134	5.8	1599	2 T15854	hypothetical prote
9	133	5.7	675	2 S60612	protein-tyrosine k
10	132.5	5.7	592	1 LBLR	actin-binding prot
11	132	5.7	1503	2 S17983	gene posterior sex
12	130.5	5.6	724	2 A38749	3-phosphatidylinos
13	130	5.6	559	2 I49444	SH3 binding protei
14	130	5.6	724	2 A38747	phosphatidylinosit
15	129	5.6	751	2 S68957	adhesive plaque pr
16	128.5	5.5	3421	1 WZBBB6	367K tegument prot
17	128	5.5	706	2 A45990	junctional sarcopl
18	128	5.5	724	2 A38748	3-phosphatidylinos
19	127	5.5	1385	2 T21706	hypothetical prote
20	125.5	5.4	1029	2 T30351	muclin-like protein
21	125	5.4	461	2 A57463	p55Prk - mouse
22	125	5.4	962	2 S58107	hypothetical WD-40
23	124.5	5.4	473	2 S25776	transforming prote
24	124.5	5.4	585	2 A46209	protein-tyrosine-p
25	124.5	5.4	593	1 JN0805	protein-tyrosine-p
26	124.5	5.4	597	1 A53593	protein-tyrosine-p
27	123.5	5.3	841	2 A43254	protein-tyrosine-p
28	123.5	5.3	1317	2 T03748	apoptosis associat
29	122.5	5.3	596	2 I38228	Shb protein - huma

30	122.5	5.3	875	2 S23760	polyphenolic adhes
31	121	5.2	485	2 T22147	hypothetical prote
32	120.5	5.2	431	2 T01557	hypothetical prote
33	120.5	5.2	1230	2 A36466	1-phosphatidylinos
34	120.5	5.2	1792	2 A57075	tensin - chicken (
35	120	5.2	559	2 B44265	ENL (translocation
36	120	5.2	509	1 F0W4GM	gag-abl polyprotel
37	119.5	5.2	595	1 A55651	protein-tyrosine k
38	119.5	5.2	1052	2 A82127	protein-tyrosine-p
39	119.5	5.2	1733	2 S27939	ribonuclease E VC2
40	119.5	5.2	1744	2 A54970	tensin - chicken
41	119.5	5.1	1744	2 A54970	tensin, cardiac mu
42	119	5.1	720	2 T26819	hypothetical prote
43	119	5.1	2187	2 T30826	nascent polypeptid
44	118.5	5.1	440	2 JC7807	Wiskott-Aldrich sy
45	118.5	5.1	657	2 A39526	probable protein k

ALIGNMENTS

RESULT 1

B56110

tyrosine phosphoprotein SLP-76 - mouse

C/Species: Mus musculus (house mouse)

C/Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999

C/Accession: B56110

R/Jackson, J.K.; Mottlo, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretz

J. Biol. Chem. 270, 7029-7032, 1995

A>Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with

A/Reference number: A56110; MUID:95221345; PMID:7706237

A/Accession: B56110

A/Status: preliminary

A/Molecule type: mRNA

A/Residues: 1-533 <JAC>

A/Cross-references: GB:U20159; NID:9806767; PIDN:AAC52189.1; PID:9806768

C/Superfamily: SH2 homology

C/Keywords: phosphoprotein

F/422-520/Domain: SH2 homology <SH2>

Query Match	17.8%	Score 412.5	DB 2	Length 533
Best Local Similarity	29.7%	Pred. No. 2.1e-20		
Matches 148	Conservative 51	Mismatches 175	Indels 125	Gaps 20
QY 14 FGLDRFQNVSL-----KNRSMFSLSAKRCRAVLPLPDHRRNLG-----	56			
DB 57 FPLRLMPLLSKLSQDINKNEERSIFRKPQIPRFLEFESHEDDGMSFEDDEYSPN 116				
QY 57 --VPGGKCNNDYEDP--EFOLL-----KAMP-----MKILPARPIQ--ESEY 96				
DB 117 DDDPDGE---DDGDYSPNEEDQALVDADADYEPSPSNEEALQSSITLPPNSFHTNSMT 173				
QY 97 ADTRYFQDMMEAPLLP--PKASY-----	118			
DB 174 IDRPPTGKVSQPPVPLPRKPPALPLPTGRNHSPLSPPHNHEBPSRSGNNKTAKLPAP 233				
QY 119 STERQRPDVMTQLEVD-----KPTKDYRSQRFKFKTKINKTLPPLPPRAITL 170				
DB 234 SIDRSTKPPLDRLAPLDREPFILGKKPPSDPSAPLGRHLPKIOKPLP---PAMDR 290				
QY 171 PKRYQPLPAPPPRESSAYFAKPTPEVORGPPORSAKDRSVLGAEEESHQTKPSSC 230				
DB 291 HENNERLGPVTR-----KPSVPHGSGPDRRE-----NDEDVHQRPPLPQPSL 334				
QY 231 PSSNQNTQKSPPAIAS-----SYMGKHSIQARDHTGSMQC--PAORCOAAASHSP- 281				
DB 335 PSMSSNTFFRSYQPSKNTFFPLAHMGARS--ESNIGQASQASLPPTYSQGPGRNPL 391				
QY 282 ----RMPIYENTSEKPDPTKPEKDY--WQNEYIGTYSQAVEDVLMKENKGTFLVRD 336				
DB 392 RSEGRNLPLEVPN--RPQSPSPGEETPLDEEYVSATIFPEAEALRLKINQGTFLVRD 449				
QY 337 CSFKSAEPYVLVVFQGNKYVNVKIRFLBSNQGFALGTGLRGNEMDSDVEDIIEHTTYP 396				

Db 450 SSKTANNPYVLWLVKXKYNINQIRYQESQVYLGTLGKNGKEDFLSVSDIIDYFRKMP 509
 397 ILLIDGKDKAARRKQCYIT 415
 Db 510 LLLIDGKNRGS-RYQCTLT 527

RESULT 2

A56110
 Tyrosine phosphoprotein SLP-76 - human
 C:Species: Homo sapiens (man)
 C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
 R:Accession: A56110
 J:Biochem. 270, 7029-7032, 1995
 A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with G
 A:Reference number: A56110; MUID:95221345; PMID:7706237
 A:Accession: A56110
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-533 <JNC>
 A:Cross-references: GB:020158; NID:9806765; PIDN:AMC50135.1; PID:9806766
 C:Gene: SLP-76
 A:Cross-references: GB:1230199; OMIM:601603
 A:Map position: 5q33.1-5qter
 C:Superfamily: SH2 homology
 C:Keywords: phosphoprotein
 F:422-520/Domain: SH2 homology <SH2>

Query Match 17.2%; Score 398; DB 2; Length 533;
 Best Local Similarity 30.2%; Pred. No. 2e-19;
 Matches 124; Conservative 52; Mismatches 146; Indels 88; Gaps 14;

Qy 68 DYEDPEQLKAMPKILPAPRIQES--YADTRYEQDMAP-----LLPPKA 116
 Db 144 DYEPNSDEEALQN-SILPAKFPNSMSYIDRPPSGKTPOQPVPPQPMALPPPA 202
 Qy 117 -----SVSTEROTRDVNRTOLEVD-----PPT 139
 Db 203 GRNHSPPLPPPTNHEPSSRNHKTAKLPASIDSTKPLDRLSLAPDREPTTGKPP 262
 Qy 140 FKD--VRSQRFKPKYKINKTPPLPPRPATTLKPKYQPLPAPPESSAYFAKPTFP 196
 Db 263 FSDKPSIFAGRSIGELHLPKLP-----TERHERSSPLP-----GKKPPY 308
 Qy 197 EVQRPGRORSADKFSRVLGAEESHQTKPE-----SSCPSSNQNTQKSPPAIASSSY 250
 Db 309 KKGWCPDRRE-----NDEDDVHQRPPLPALLPMSSNTPFSRSTKPSMNPPLPSSHM 360
 Qy 251 PGKHSIARDHTGSMQHCPCAPCOAASHSP-----RMLPYENTSEKPDPTKPEKDW 305
 Db 361 PEAFF-ESSNSFPQASASLPYFSGSPSNPPRAAGRNPLPLPKKPR-PSPAEEENSL 418
 Qy 306 QNEWYIGEXSROAVEDVLKKNKDGTFVRCSTKSAKEPVYLVYFYGNKYVNVIRPLE 365
 Db 419 NEEWVSYITREAEALRKINQDCTFLVROSSKTTNPYLVWLYKDKVYNIDIRYQK 478
 Qy 366 SNOQALGTGLGNGEMFDSVEDIIEHYTFPILLIDGDKAARRKQCYIT 415
 Db 479 ESQVYLVLTGLGKEDFLSVSDIIDYFRKMPULLIDGKNRGS-RYQCTLT 527

RESULT 3

TWVNR
 Protein-tyrosine kinase (EC 2.7.1.112) fgr - feline sarcoma virus (strain Gardner-Rashee
 C:Species: feline sarcoma virus
 A:Note: host Felis sp. (cat)
 C:Date: 27-Nov-1985 #sequence_revision 26-May-1995 #text_change 31-Mar-2000
 C:Accession: A00653; A03937
 R:Naharro, G.; Robbins, K.C.; Reddy, E.P.
 Science 223, 63-66, 1984

A:Title: Gene product of v-fgr onc: hybrid protein containing a portion of actin and
 A:Reference number: A00653; MUID:84097512; PMID:6318314

A:Accession: A00653
 A:Molecule type: DNA
 A:Residues: 1-663 <NAH>
 A:Cross-references: GB:X00255; GB:K01487; NID:961542; PIDN:CAA25063.1; PID:961543
 A:Note: the authors translated the codon GAT for residue 14 as Glu
 A:Comment: This protein is synthesized as a gag-fgr polypeptide.
 C:Gene: fgr
 C:Superfamily: feline sarcoma virus protein-tyrosine kinase fgr; protein kinase homol
 C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; p
 F:1-118/Region: gag polypeptide similarity
 F:141-268/Region: actin similarity
 F:285-382/Domain: SH2 homology <SH2>
 F:402-660/Domain: protein kinase homology <KIN>
 F:410-418/Region: protein kinase ATP-binding motif
 F:432/Active site: Lys #status predicted

Query Match 6.7%; Score 154.5; DB 1; Length 663;
 Best Local Similarity 23.4%; Pred. No. 0.007;
 Matches 71; Conservative 34; Mismatches 91; Indels 107; Gaps 15;

Qy 161 LPPRPATTLPRKY-----QPLPAP-----EESAYF----- 189
 Db 101 LPPRPATSLQPHSROPALCRPVCRRPLPLPTAMEEYVALVINDSGMCKAG 160
 Qy 190 ----APKPTPEYVGRGRSA-----KQSRVLGAEESHQTKRESSCPSSNQNTQ 238
 Db 161 FAGDDAPRAVPFISVGRHOGVWYGQKD--SYVGDAQSKR----- 202
 Qy 239 KSPPAIASSSYWPKGHSIOAR--DHTGSMQHCPCAPCOAASHSPRMLPYENTSEKPD 295
 Db 203 ---GILTKY-PIEHGIVTMNDMEKIMHFTYELNLVAPEEHVLL-----TEAVLN 251
 Qy 236 PTKPEK-----DYQNEWYIGEXSROAV-DVLKKNKDGTFV 334
 Db 252 PRANREKMTQINFETFNIPSNVAVPVDISOAEWYFGKIGRDAEROLLSGNAGARLV 311
 Qy 335 RDCSTKSAKEPVYLVV-----FYGNKYVNVIRFESNQQFALGTGLGNGEMFDSVEDII 389
 Db 312 RESETTKA--YSLIRWDEAKGDVAKYIKRLKIDG-----GYITTRAFQNSVQELV 364
 Qy 390 EHY 392
 Db 365 QHY 367

RESULT 4

phosphoinositide-3-kinase regulatory beta chain [imported] - human

A:Accession: H59435
 C:Species: Homo sapiens (man)
 C:Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
 C:Accession: H59435; A59436
 R:Volinia S; Patrachini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L
 Oncogene 7, 789-793, 1992
 A:Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinosit
 A:Reference number: H59435
 A:Accession: H59435
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <VOL>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1
 R:Janssen, J.W.; Schleithoff, L.; Bartram, C.R.; Schulz, A.S.
 Oncogene 16, 1767-1772, 1998
 A:Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta sub
 A:Reference number: A59436; MUID:98241181; PMID:9582025
 A:Accession: A59436
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <JAN>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1

Query Match 6.3%; Score 146; DB 2; Length 728;
Best Local Similarity 22.9%; Pred. NO. 0.03;
Matches 87; Conservative 48; Mismatches 147; Indels 98; Gaps 18;

[illegible]

RESULT 5

Cell 65, 91-104, 1991
A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine kinase
A:Reference number: A38749; MUID:91191567; PMID:1707345
A:Accession: B38749
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-723 <ORS>
A:Cross-references: GR:M61745; GB:M61746
C:Superfamily: SH2 homology
C:Keywords: phosphotransferase
F:345-420/Domain: SH2 homology <SH2A>
F:617-706/Domain: SH2 homology <SH2>

Query Match	6.2%	Score 142.5;	DB 2;	Length 723;
Best Local Similarity	25.4%;	Pred. No. 0.051;		
Matches 61;	Conservative 23;	Mismatches 79;	Indels 77;	Gaps 11

[illegible]

RESULT 6
T27877

hypothetical protein ZK470.5 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T27877
 R:Minx, P.
 submitted to the EMBL Data Library, October 1995
 A:Description: The sequence of C. elegans cosmid ZK470.
 A:Reference number: Z20433
 A:Accession: T27877
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-443 <MIN>
 A:Cross-references: EMBL:U39651; PIDN:AAA80397.1; CESP:ZK470.5
 C:Genetics:
 A:Gene: CESP:ZK470.5
 A:Introns: 39/3; 83/3; 109/3; 214/3; 270/3; 332/3; 352/3; 416/2

Query Match	6.0%;	Score 138.5;	DB 2;	Length 443;
Best Local Similarity	28.6%;	Pred. No. 0.051;		
Matches 46;	Conservative 23;	Mismatches 49;	Indels 43;	Gaps 8;

[illegible]

RESULT 7

Shb-like adapter protein, Shf - human
C:Species: Homo sapiens (man)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: J07552
R:Lindholm, C. K.; Frantz, J. D.; Shoelson, S. E.; Welsh, M.
Biochem. Biophys. Res. Commun. 278, 537-543, 2000
A:Title: Shf, a Shb-like adapter protein, is involved in PDGF-alpha-receptor regulation
A:Reference number: J07552; MUID: 20548990; PMID:1109546
A:Accession: J07552
A:Molecule type: mRNA
A:Residues: 1-480 <LIN>
C:Comment: This protein, a novel adapter protein, has the roles in PDGF-receptor and
C:Genetics:
A:Gene: Shf
C:Keywords: apoptosis

Query Match	5.88	Score 134.5	DB 2	Length 480
Best Local Similarity	20.78	Pred. No. 0.11		
Matches 85	Conservative 55	Mismatches 138	Indels 133	Gaps 19

```
QY      26 KNRSMPLSSAK -GRCAVLELP-----DHRNLAVPGG- EKONSNDY   69
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     .135 KKSNRSTTRLRIRLDRLEPRRLAILEDYADFEVDQEGGSAGSGAPEKPENDGY 194
QY      70 EDPEQLLKAMPMSKITPARIQESEYADTFRFODMMEAPLL-----APKASVSTERO 133
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     195 MEPR-----YEAOKM-----AEIRGSKETTLQPLPYDIPIEEDBGATPEGE 237
QY      124 -----TRDVPMQTQ-----LEEWDKP-TFKDVNSORFKEFTYKTINKTLPMP-----RPA 167
        | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     238 GAPMRPREERLPDEDERPREEIDQWEMKKKEISAFNADIKVKIDLPMPPPVQOLDSFS 297
QY      168 I-----TLPKKYQPLPAPPEESSAYRAPKTPFEVORGPRORSAKDFSVLGAEEESH 222
```

Db 218 LPDGRDIDSGPASLPPESELEDSASOF----- 324
 Oy 223 QTKPSCPCSSNQNTQ-KSPPAIASSSYPMGKHSTQARDHTGSMQHCPRACQAAASHSP 281
 C:Species: *Caenorhabditis elegans*
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
 C:Accession: T15854
 R: Fulton, L.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of *C. elegans* cosmid C56C10.
 A:Reference number: 218417
 A:Accession: T15854
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1599 <FULL>
 A:Cross-references: EMBL:U29488; NID:9868238; PID:9868250; PIDN:AAA68780.1; CESP:C56C10.
 C:Genetics:
 A:Experimental source: Strain Bristol N2
 C:Gene: CESP:C56C10.12
 A:Interons: 38/3; 59/2; 92/3; 178/3; 380/1; 424/3; 554/1; 573/3; 599/2; 697/1; 738/3; 980
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein C56C10.12

RESULT 8

T15854
 hypothetical protein C56C10.12 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 24-Nov-1999
 C:Accession: T15854
 R: Fulton, L.
 submitted to the EMBL Data Library, June 1995
 A:Description: The sequence of *C. elegans* cosmid C56C10.
 A:Reference number: 218417
 A:Accession: T15854
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-1599 <FULL>
 A:Cross-references: EMBL:U29488; NID:9868238; PID:9868250; PIDN:AAA68780.1; CESP:C56C10.
 C:Genetics:
 A:Experimental source: Strain Bristol N2
 C:Gene: CESP:C56C10.12
 A:Interons: 38/3; 59/2; 92/3; 178/3; 380/1; 424/3; 554/1; 573/3; 599/2; 697/1; 738/3; 980
 C:Superfamily: *Caenorhabditis elegans* hypothetical protein C56C10.12

Query Match 5.8%; Score 134; DB 2; Length 1599;
 Best Local Similarity 23.7%; Pred. No. 0.51;

Matches 56; Conservative 34; Mismatches 82; Indels 64; Gaps 13;

Oy 85 ILPAPDGESEVADRYQDDMEAP-----LLLP-----PKASVSTRQTRDVRM 129
 Db 5 VPKKP-KHREKQSDYPPIDPRALVNGRLPAASHTIEDLERSASPEEDQDIST 63
 Oy 130 TQ---LEVDKPTFKDYRSQRFKGYTKINKTLPPLPRAITLPKKYQPLP----- 179
 Db 64 TEDSLKREASBP-LKDYRSSVR-----PAPP-PRVSEKERAPPIPPSMILFP 111
 Oy 180 -----APPESSAYFAPK-----PTPEVQGRFQRSKADFSRYL---GAEEES 220
 Db 112 RSTSMVAESRKSTAVAPKRSVAVASYPAYVELAELPSYDALQHPQVYPSINGLQHS 171
 Oy 221 HHQT-----KPESSCPSSNQNTQ-KSPPAIASSSYPMGKHSTQARDHTGSMQHCPR 269
 Db 172 HSATAIPKTRFSAVEREREREGCAPMPYSIKTYERNEHGLMTEMLVTFYHNP 227

RESULT 9

S60612
 protein-tyrosine kinase (EC 2.7.1.112) Bmx - human
 C:Species: *Homo sapiens* (man)
 C:Date: 27-Apr-1996 #sequence_revision 13-Mar-1997 #text_change 18-Feb-2000
 C:Accession: S60612
 R: Tamagone, L.; Lahtinen, I.; Mustonen, T.; Vitaneva, K.; Francis, F.; Muscatelli, F.;
 Oncogene 9, 3683-3688, 1994
 A:Title: Bmx, a novel nonreceptor tyrosine kinase gene of the BTK/ITK/TEC/TYK family loc
 A:Reference number: S60612; MUID:95060827; PMID:7970727
 A:Accession: S60612
 A:Status: preliminary; nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-675 <TRM>
 A:Cross-references: EMBL:X83107; NID:9951234; PIDN:CAA58169.1; PID:9951235

C:Genetics:
 A:Gene: Bmx
 C:Superfamily: protein-tyrosine kinase tec; pleckstrin repeat homology; protein kinas
 C:Keywords: ATP; phosphotransferase; tyrosine-specific protein kinase
 F:3-109/Domain: pleckstrin repeat homology <PLK>
 F:296-392/Domain: SH2 homology <SH2>
 F:415-673/Domain: protein kinase homology <KIN>
 F:423-431/Region: protein kinase ATP-binding motif

Query Match 5.7%; Score 133; DB 2; Length 675;
 Best Local Similarity 24.0%; Pred. No. 0.2;

Matches 62; Conservative 41; Mismatches 107; Indels 48; Gaps 11;

Oy 166 PAITLPKTKQPLPPAPPESSAYFAPKPPFC-VORGP-----QRSKDFSRVGA--- 216
 Db 141 PCCTLMENAVANLHTVNEKRV-----PTPEPVLRPAVAVLKMDFSSSTTLAQTDN 196
 Oy 217 EESHQTKPESSCPSSNQNTQKSPPAIASS-----SYWPG-----KHSIQAR 259
 Db 197 ESKNYGSPPESSSTSLAQYDSNSKKYGSQNFMMQYIPREDPDMQVRLKSSSSE 256
 Oy 260 DHTGSMQHCPRACQAAASHSPMLPYENTNSEKPDPTKPEKDYWQNEWTGYEQAV 319
 Db 257 DVASSNQ-----BRNVHTTSKISWEPSSSE---EENLDYDWFAGNISRSQS 306
 Oy 320 EDVLKKNKDGTFELVDCSTKSKAPPYLVF---YGNKYVWKIRPLESNQOALGTGL 376
 Db 307 EDLLQKQKREGAFMYRN---SSQVGMVYSLFSKAVNDKKGVKHYHTNAENKL--YL 361
 Oy 377 KQNEPDSVEDIIEHYTY 394
 Db 362 AENYCFDSIPKLIHYHOH 379

RESULT 10

LIBY
 actin-binding protein ABP1 - yeast (*Saccharomyces cerevisiae*)

N:Alternate names: protein YCR088w
 C:Species: *Saccharomyces cerevisiae*

C:Date: 31-Mar-1991 #sequence_revision 31-Dec-1993 #text_change 05-Nov-1999
 C:Accession: S19503; S19767; S07608

R: Dusterhoft, A.; Erdmann, D.; Hegemann, J.; Philippsen, P.; Schweitzer, B.; Spiegelb
 submitted to the Protein Sequence Database, March 1992

A:Reference number: S19351
 A:Accession: S19503

A:Molecule type: DNA
 A:Residues: 1-476 <DUS>

A:Cross-references: EMBL:X59720; GSPDB:GN00003; MIPS:YCR088w
 R:Fontall, L.; Grisanti, P.

submitted to the Protein Sequence Database, March 1992
 A:Reference number: S19504
 A:Accession: S19767

A:Molecule type: DNA
 A:Residues: 14-592 <PRO>

A:Cross-references: EMBL:X59720; GSPDB:GN00003; MIPS:YCR088w
 R:Drubin, D.G.; Mulholland, J.; Zhu, Z.; Botstein, D.

Nature 343, 288-290, 1990
 A:Title: Homology of a yeast actin-binding protein to signal transduction proteins an

A:Reference number: S07608; MUID:90136906; PMID:2405273
 A:Accession: S07608

A:Molecule type: DNA
 A:Residues: 1-57,'S',59-311,'I',313-592 <DRU>

A:Cross-references: EMBL:X51780; NID:933321; PIDN:CAA36075.1; PID:933322
 C:Genetics:

A:Gene: SGD:ABP1; MIPS:YCR088w
 A:Cross-references: SGD:S0000684; MIPS:YCR088w

A:Map position: 3R
 C:Superfamily: actin-binding protein ABP1; SH3 homology

C:Keywords: actin binding; duplication
 F:158-180,378-401/Region: duplication

F:339-388/Domain: SH3 homology <SH3>
 Query Match 5.7%; Score 132.5; DB 1; Length 592;

Science 259, 1157-1161, 1993
 A>Title: Identification of a ten-amino acid proline-rich SH3 binding site.
 A:Reference number: 149444; MUID:93174278; PMID:8438166
 A:Accession: 149444
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-559 <RES>
 A:Cross-references: GB:L14543; NID:9293267; PIDN:AAA37121.1; PID:9293268
 C:Superfamily: pleckstrin repeat homology
 F:25-128/Domain: pleckstrin repeat homology <PLK>
 F:201-210/Region: proline-rich SH3 binding

Query Match 5.6%; Score 130; DB 2; Length 559;
 Best Local Similarity 21.4%; Pred. No. 0.26;
 Matches 87; Conservative 42; Mismatches 145; Indels 132; Gaps 19;

OY 86 LPAPRQSEY---ADPRYDMMEARLL-----PPKASSTEGOTRDVMT 130
 DB 16 LSSYFDNEDYEHEDDDSYLPPSPGPKLEDALTYPPAPPPVPR----- 213
 OY 131 QLEVDKPTFKDVSQRFKFKYTKINKTP-LPPRPATILP-----KKY 174
 DB 214 -----KPAFSDLR-----RAHSFTSKSPRLPPRPKRGLPDTGSAPEDAKDALGLRV 263
 OY 175 QP---LPAPRPESSAIFYAPKFTPEVQGRORS-----A 207
 DB 264 EPGLRVPATPRRMSDPMPMSNPTVPLNRKPCFRDSVNPGLPMTPGHGTSSVSTMA 323
 OY 208 KDFSVLGAEEESHQTK--PESSCPSSNQNTQ-----SPPAISSSTMPCKHSIQ 257
 DB 324 VATSNCKDKLKFLLSSKPPSEPPVPANKPKFLKAEPSPREAKFAVP--PVA 380
 OY 258 ARDHTGSHQCPCAPQCAASHSPMLPYENTNSEKPD-----PTKPD--- 300
 DB 381 PRPPQKMP-MPEATVRAVLPREPNTPLPHLQSPRPGQSRGSEFKARQPSADG 439
 OY 301 -----EKDWQONWTYIGESRQAVEDVLMK-----ENKDGTFIVRCCSTKSKAPVYL 348
 DB 440 EDSDEYKVKPLPNSVFNTTESCVERLKFATDPRGEPQGLCYCIRNSFTKSGK--VL 496
 OY 349 VFE--YGNKYVNVKIRPLESNQOAFGLGNGEMEDVEDIIEHY 392
 DB 497 VVWDESSKVKVNRKI--FEKDSKF-----LREGEVLFASVGSVMEY 536

RESULT 14
 A38747
 C:Species: Mus musculus (house mouse)
 C:Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 16-Jul-1999
 C:Accession: A38747
 R:Escobedo, J.A.; Navankasattusas, S.; Kavanaugh, W.M.; Milfay, D.; Fried, V.A.; William
 Cell 65, 75-82, 1991
 A>Title: cDNA cloning of a novel 85 kd protein that has SH2 domains and regulates bindin
 A:Reference number: A38747; MUID:91191564; PMID:1849460
 A:Accession: A38747
 A:Molecule type: mRNA
 A:Residues: 1-724 <ESC>
 A:Cross-references: GB:M60651
 C:Comment: This protein binds a phosphotyrosine-containing sequence of ligand-activated
 phosphatidylinositol at position 3 of the inositol ring.
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:33-42/Domain: SH2 homology <SH2A>
 F:624-718/Domain: SH2 homology <SH2>

Query Match 5.6%; Score 130; DB 2; Length 724;
 Best Local Similarity 22.0%; Pred. No. 0.35;
 Matches 89; Conservative 46; Mismatches 147; Indels 122; Gaps 19;

OY 38 GRCAVLEPLDHR--RLTAGVPGCKNSNND--PEFOLKAWPSMKILPAPRQES 94
 DB 78 GRKR-ISPPTPKRPPRPLPVAPOSSKTEADTEQALPLPLAEQAFAPDPVAPLLIKLL 136

OY 95 EXADTRYDMMEARLLPPKASVSTERQTR-----DYRMTQLEVDKPTFKDVSQRFK 149
 DB 137 EAIEKK-----GLECSTLYRTQSS--SNPAELRLDCCDAASVDLEMDIVLADA---FK 187
 OY 150 GFKYTKINKTPLP-----PPRATILPKKYQPLPPAPPE-ESSAYFAPK 192
 DB 188 --RYLADLPNPVITPVAVYVNEWMSLAQELQSPEDCQLKLKILRLNIRHQCLTLQIYLK 245
 OY 193 PTFPEVQGRORSKAD-----FSRVL-----GAEEESHQTPRESSCPSSNONT 237
 DB 246 HFEK-----LSQASSKNILNARVLSEIFSPVLFREPAASDNTHELKAIETLISTENNE 300
 OY 238 OKSPPAISSSTMPCKHSIQARDHTGSHQCPCAPQCAASHSPMLPYENTNSEKPPPT 297
 DB 301 KOPAPAL-----PPKPP 312
 OY 298 KP-----DEKDWONWTYIGESRQAVEDVLMKENKDGTFIVRCCSTKSKAPVYL 348
 DB 313 KPTVANNMNNMNSLQDAEWYWDISREVENKL--RDTAGTFIVRDASTKMHGD-YTL 370
 OY 349 VVFYGNKYVNVKIRPLESNQOAFGLGNGEMEDVEDIIEHY 392
 DB 371 TLKRGGN--NKLKIKIFHRDGKYGFSDDL-----TFNSVVELINHY 408

RESULT 15
 S68957
 C:Species: Mytilus galloprovincialis (Mediterranean mussel)
 C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 18-Feb-2000
 C:Accession: S68957
 R:Inoue, K.; Takeuchi, Y.; Miki, D.; Odo, S.; HaraYama, S.; Waite, J.H.
 Submitted to the EMBL Data Library, August 1995
 A:Description: Cloning, sequencing and sites of expression of genes for the hydroxylar
 A:Reference number: S68957
 A:Accession: S68957
 A:Molecule type: mRNA
 A:Residues: 1-751 <INOC>
 A:Cross-references: EMBL:D63778; NID:9961463; PIDN:BA00981.1; PID:01010497; PID:9961
 C:Keywords: hydroxyproline
 F:1-24/Domain: signal sequence #status predicted <SIG>
 F:25-751/Product: adhesive plaque protein #status predicted <MAT>
 F:123,127,133,173,177,183,227,267,283,293,297,303,307,313,323,327,333,337,343,347,367
 33/Modified site: 4-hydroxyproline (Pro) #status predicted
 F:125,129,135,139,175,179,185,189,225,229,265,269,285,289,295,299,305,309,315,319,325
 79/Modified site: 3',4'-dihydroxyphenylalanine (Tyr) #status predicted
 F:126,136,176,186,226,266,286,296,306,316,326,336,346,366,376,386,406,416,426,436,446
 ans-2,3-cis-3,4-dihydroxyproline (Pro) #status predicted
 F:485,489,495,499,505,509,515,519,525,529,535,539,545,549,555,559,565,569,575,579,585
 dihydroxyphenylalanine (Tyr) #status predicted
 F:537,543,547,553,567,573,577,583,587,593,597,603,607,613,617,623,627,633,637,643,647

Query Match 5.6%; Score 129; DB 2; Length 751;
 Best Local Similarity 23.4%; Pred. No. 0.43;
 Matches 55; Conservative 34; Mismatches 108; Indels 38; Gaps 9;

OY 84 KILARPQSEYVADTR---FQDMMEARLLPPKASVSTERQTRQRMQDLEVDKPTF 140
 DB 88 KLSSTYKPLKTYNNKNTNPPVYKPKMTYPTYYKPKSPPTYYK-----PPSY 135
 OY 141 KDVSORFKGFKYTKINKTP-LPPRPATILPDKKYQPLPPAPPESSAYFAKPPFPEVQR 200
 DB 136 PATYKSSSYSSYKPKPKTYPTYYKPLTYPTYYKPKPSYPT-----YKPKPYPATYK 190
 OY 201 G-----PRQSAKDFSRVLAEEESHQTKPESSCPSSNONTQKSP-----AIASSSYM 250
 DB 191 SKSSYPSYKTKKTYPSYKPKTYPTYYKPKSYPTYYKSKKSYPTYYTKASYPSYK 250
 OY 251 PGKHSIQARDHTGSHQCPCAPQCAASHSPMLP-YENTNSEKPD--PT---KP 299
 DB 251 PKK--TYPSTYKPKISTYPTYYKAKPSYPTSYRAKPSYPTYYKAKPSYPTYYKAP 303

Tue Apr 22 16:29:16 2003

us-09-856-061-2.rpr

Page 7

Search completed: April 21, 2003, 12:38:34
Job time : 20.6276 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:28:57 ; Search time 9.65475 Seconds

(without alignments)
1868.739 Million cell updates/sec

Title: US-09-856-061-2

Sequence: 1 MTSQGNKRTKEGFGDLRFQ.....QPLPLARLLITQYSSQALHE 435

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database: SwisProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	412.5	17.8	533	1 LCP2_MOUSE	Q60787 mus musculu
2	398	17.2	533	1 LCP2_HUMAN	Q13094 homo sapien
3	146	6.3	724	1 P85B_BOVIN	P23726 bos tauris
4	146	6.3	724	1 P85B_HUMAN	O00459 homo sapien
5	143.5	6.2	722	1 P85B_RAT	O63788 rattus norv
6	142	6.1	545	1 FGR_FSVGR	P00544 feline sarc
7	140	6.0	540	1 SCK_HUMAN	P98077 homo sapien
8	136.5	5.9	561	1 3BP2_HUMAN	P78314 homo sapien
9	133	5.7	675	1 BMX_HUMAN	P51813 homo sapien
10	133	5.7	724	1 P85A_RAT	O63787 rattus norv
11	132.5	5.7	217	1 GRAP_HUMAN	O13588 homo sapien
12	132.5	5.7	271	1 CRK_DROME	O94900 drosophila
13	132.5	5.7	592	1 ABL1_YEAST	P15891 saccharomyc
14	132	5.7	1603	1 PSC_DROME	P35820 drosophila
15	130.5	5.6	722	1 P85B_MOUSE	O08908 mus musculu
16	130.5	5.6	724	1 P85A_BOVIN	P23727 bos tauris
17	130	5.6	559	1 3BP2_MOUSE	O06649 mus musculu
18	130	5.6	724	1 P85A_MOUSE	P26450 mus musculu
19	129	5.6	461	1 P55G_HUMAN	O92569 homo sapien
20	129	5.6	751	1 FPL_MYTG	O27409 mytilus gal
21	128.5	5.5	3421	1 TEGU_HSVB	P28605 teguine herp
22	128	5.5	705	1 TRDN_RABIT	O28820 corytaeagus
23	128	5.5	724	1 P85A_HUMAN	P27986 homo sapien
24	127	5.5	461	1 P55G_BOVIN	O46404 bos tauris
25	127	5.5	1362	1 BRD4_HUMAN	O60885 homo sapien
26	125	5.4	461	1 P55G_MOUSE	O64143 mus musculu
27	125	5.4	962	1 YAE_SCHPO	O09731 schizosacch
28	124.5	5.4	583	1 SRC_HUMAN	P29353 homo sapien
29	124.5	5.4	585	1 PTNB_MOUSE	P35235 mus musculu
30	124.5	5.4	593	1 PTNB_HUMAN	Q06124 homo sapien
31	123.5	5.3	845	1 CSW_DROME	P29349 drosophila
32	123.5	5.3	1386	1 ZAP3_MOUSE	O9017 mus musculu

34	122.5	5.3	875	1 FPL_MYTED	O25460 mytilus edu
35	120.5	5.2	1290	1 ENL_HUMAN	P19174 homo sapien
36	120	5.2	559	1 ENL_HUMAN	O03111 homo sapien
37	119.5	5.2	509	1 STK_HYDAT	P17713 hydra atten
38	119.5	5.2	1490	1 CRK7_HUMAN	O94904 homo sapien
39	119.5	5.2	1744	1 TENS_CHICK	O04205 gallus gall
40	119	5.1	517	1 FGR_MOUSE	P14234 mus musculu
41	118.5	5.1	526	1 SRC_RSVSR	P00524 rous sarcom
42	118.5	5.1	657	1 SGV1_YEAST	P23293 saccharomyc
43	118.5	5.1	659	1 BTK_HUMAN	O06187 homo sapien
44	118	5.1	2774	1 MAPA_RAT	P34926 rattus norv
45	117.5	5.1	659	1 BTK_MOUSE	P35991 mus musculu

ALIGNMENTS

RESULT 1
LCP2_MOUSE STANDARD: PRT: 533 AA.
AC Q60787;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
GN LCP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell lymphoma;
RX MEDLINE=95221345; PubMed=7706237;
RA Jackman J.K., Motta D.G., Sun Q., Tanemoto M., Turk C.W., Peltz G.A., Koretzky G.A., Finkel P.R.;
RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with Grb2 in T cells";
RT J. Biol. Chem. 270:7029-7032(1995).
RL
CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND FVB.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES.
CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION (BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

CC EMBL: U20159; AAC52189.1; -;
DR HSSP: P12931; ISHD.
DR MGD: MGI:1321402; LCP2.
DR InterPro: IPR001660; SAM.
DR InterPro: IPR00980; SH2.
DR Pfam: PF00017; SH2; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00454; SAM; 1.
DR SMART: SM00252; SH2; 1.
DR PROSITE: PS00001; SH2; 1.
KW SH2 domain; Phosphorylation.
FT DOMAIN 422 530
SQ SEQUENCE 533 AA: 60228 MW: 60DB77782C2E60C3 CRC64:
Query Match 17.8%; Score 412.5; DB 1; Length 533;
Best Local Similarity 29.7%; Pred. No. 1.7e-19;

Matches 148; Conservative 51; Mismatches 175; Indels 125; Gaps 20;

QY 14 FGDRLFGQVSL-----KNRSMPSLSAKGRCRAVLPLPHRRRLAG-----56
 DB 57 FPKLMLPLSLKSDINKNEERRSTFTTKPPIPRPLETESHEDDGWSSFEEDYESPN 116
 QY 57 --VEGEGECNSNDYEDP--EFQLL-----KAMP-----MKILPAPPIQ--ESEY 96
 DB 117 DDDPGE---DDGYESEFNEEOALVDDADYEPSPNNEBALQSLIPSPFNNTSMY 173
 QY 97 ADTRFYQDMMEAPLLP--PKASY-----118
 DB 174 IDRPRTGVSGQPPVPPRLKPKAPLPLEPTGNNHSLPSPHPHNEPSPSGNNKTAKLPAP 233
 QY 119 STERQTRDVRMTQLEVD-----KPTFKDVSORFKGKRYKINKTPPLPPRAPITL 170
 DB 234 SIDRSTKPLDRLSLAPLDREPFILGKKRPFSDKPSAPLGRHRLKIQKPLP---PAMDR 290
 QY 171 PKKYOPLPAPPESSAVFAKPFPEVQGRORSAKDSRYLGAEEESHQTPRESSC 230
 DB 291 HERNERLGPVTR-----KPSVPRHGRCPDRRE-----NDEDDVHORPLPQPSL 334
 QY 231 PSSNQNTOKSPPAIASS-----SYMGKHSIQARDHTGSMQC--PAQCOAAASHSP- 281
 DB 335 PMSNSTPFSRSVQPSKNTFPLAHMGAFS---ESNIGQOQASLPPYTSQCGGNAPPL 391
 QY 282 ----RMLPYENTNSEKPDPTKPEKDV--WQNEWYIGESQAVEDVLMKENKDTFLVRD 336
 DB 392 RSEGRNLPPLVPN--RQPPSPGGEETPLDEEYVSYTRPEAAALRKINQDCTFLVRD 449
 QY 337 CSTSKRKEPYLVVYFGNKVYKIRFLESNOQFALCTGRCNMFESVEDIIEHYTFP 336
 DB 450 SSKRTANPPLVLMVLDKYNIOIRYQESQVYLLGTGLRGKEDFLSVSDIIDYFRKMP 509
 QY 397 TLIDGKDKARRKQCYLT 415
 DB 510 LLLIDGKNRGS-RYQCYLT 527

RESULT 2
 LCP2_HUMAN STANDARD; PRT; 533 AA.
 ID LCP2_HUMAN
 AC Q13094;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
 GN LCP2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Leukemia;
 RX MEDLINE=95221345; PubMed=7706237;
 RA Jackman J.K., Motto D.G., Sun Q., Tanemoto M., Turk C.W., Peltz G.A., Koretzky J.G., Finkel P.R.,
 RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with Grb2 in T cells."
 RL J. Biol. Chem. 270:7029-7032(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=prostate;
 RA Strusberg R.;
 RL Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
 CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND F1B.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES. HIGHLY EXPRESSED ALSO IN T CELL AND MONOCYTIC CELL LINES, EXPRESSED AT LOWER LEVEL IN B CELL LINES.

CC NOT DETECTED IN FIBROBLAST OR NEUROBLASTOMA CELL LINES.
 CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION BY ZAP-70.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- DATABASE: NAME=PRO; NOTE=PRO 1:1-5(2000);
 CC WWW=ftp://www.ncbi.nlm.nih.gov/prov/guide/1118450040.g.htm".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U20158; AAC50135.1; -
 DR EMBL: BC016618; AAH16618.1; -
 DR HSSP: P12931; ISH0.
 DR Genew: HGNC:6529; LCP2.
 DR MIM: 601603; -
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2.1.
 DR ProDom: PD000093; SH2.1.
 DR SMART: SM00454; SAM; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2.1.
 DR SH2 domain; Phosphorylation.
 FT DOMAIN 133 136 POLY-GLU.
 FT DOMAIN 198 201 POLY-PRO.
 FT DOMAIN 422 530 SH2.
 SQ SEQUENCE 533 AA; 60188 MW; C5D22F31D36200C8 CRC64;

Query Match 17.2%; Score 398; DB 1; Length 533;
 Best local similarity 30.2%; Pred. No. 1.5e-18; Mismatches 124; Conservative 52; Mismatches 148; Indels 88; Gaps 14;

QY 68 DYEDPEFQLKAMPKMKIIPARPIQSE--YADTRYFQDMMEAP-----LLPPKA 116
 DB 144 DYEPSPNDEALQN-SILPAKFPNSNMYIDRPSGKTPOQPPVPPQRPMAALPPPP 202
 QY 117 -----SYTEROTRYMTQLEVD-----KPT 139
 DB 203 GRNHSLPPTQNTNHEPSSRNHKTAKLPAPSIDRSTKPLDRLSLAPFREPFILGKKPP 262
 QY 140 FKD---VRSORFGFYTKINKTPPLPPRAPITLPPKYQPLPAPPESSAVFAKPTP 196
 DB 263 FSDKPSIPAGRSIGELPLRIQKPLP-----THERHRSSPLP-----GKKPPV 308
 QY 197 EYQGRQRSADFSRVNLGAEEESHQTKPE-----SSCPSSNONTOKSPPAIASSSYM 250
 DB 309 KKGWGPDRRE-----NDEDDVHORPLPQPALPMSSTNTPSPSRSTKPSPNPLPSSIM 360
 QY 251 PGKHSIQARDHTGSMQC-PAQCOAAASHSP-----RMLPYENTNSEKPDPTKPEKDVW 305
 DB 361 PPAFS-ESNSSFPPQASLPPYTSQCGSPNRPRAAGRNFPPLPKKPPR-PSPAEEENSL 418
 QY 306 QNEWYIGESQAVEDVLMKENKDGTFVRCSTSKAKAPYLVVYFGNKVYKIRFLE 365
 DB 419 NNEWYVSYTRPEAAALRKINQDGTFLVRDSSKTTNPVLMVLDKYNIOIRYQK 478
 QY 366 SNOQFALCTGLGNEMFESVEDIIEHYTFPLIDGKDKARRKQCYLT 415
 DB 479 ESQVYLLGTGLRGKEDFLSVSDIIDYFRKMPLLIDGKNRGS-RYQCYLT 527

RESULT 3
 P85B_BOVIN STANDARD; PRT; 724 AA.
 ID P85B_BOVIN
 AC P23726;
 DT 01-NOV-1991 (Rel. 20, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase

DE p85-beta subunit) (Ptdins-3-kinase p85-beta).
GN PIK3R2.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91191567; PubMed=1707345;
RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Larrea F., Panayotou G.,
RA Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,
RA Coutnidge S.A., Parker P.J., Waterfield M.D.;
RT "Characterization of two 85 kd proteins that associate with receptor
RT tyrosine kinases, middle-T/pp60-src complexes, and p13-kinase.";
RL Cell 65:91-104(1991).
RN [2]
RP CIRCULAR DICHOISM AND FLUORESCENCE SPECTROSCOPY.
RX MEDLINE=93049176; PubMed=1330535;
RA Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski B., Dhand R.,
RA Fry M.J., Blundell T.L., Wolmer A., Waterfield M.D.;
RT Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
RT domain with a PDGF receptor phosphorylation site: structural features
RT and analysis of conformational changes.";
RL EMBO J. 11:4261-4272(1992).
CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: M61746; AAA79510.1; -
DR HSSP: P23727; 2PNB.
DR InterPro: IPR001720; PI3kinase_P85.
DR InterPro: IPR000198; RhogAP.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF000017; SH2; 2.
DR Pfam: PF000018; SH3; 1.
DR Pfam: PF00620; RhogAP; 1.
DR PRINTS: PR00678; PI3KINASEP85.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRODOM: PD000066; SH3; 1.
DR PRODOM: PD000093; SH2; 2.
DR SMART: SM00324; RhogAP; 1.
DR SMART: SM00252; SH2; 2.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50001; SH2; 2.
DR PROSITE: PS50002; SH3; 1.
KM SH3 domain; SH2 domain: Repeat.
FT DOMAIN 4 80
FT DOMAIN 125 255 RHO-GAP.
FT DOMAIN 326 421 SH2 1.
FT DOMAIN 618 712 SH2 2.
SQ SEQUENCE 724 AA; 81059 MW; 9D2BA8B6DB087098 CRC64;
Query Match 6.38; Score 146; DB 1; Length 724;
Best Local Similarity 22.38; Pred. No. 0.026;
Matches 85; Conservative 44; Mismatches 148; Indels 104; Gaps 17;
QY 46 PLPDRRLAGVPGGECNSNNDYEDPE-----FOLLKAMPKMLLPARP----- 90

DB 95 PLPARRODPPGLLPLPLPEQFSPDDVAPILVLYVAIBRTGLDSYRPPPAVRTDW 154
QY 91 -IQSEYATRTFQD-----MMEARLLPRKASSTSTORQRDVRLMDLEEVDPPTFND 142
DB 155 SLSDVEQWMAALSDGVKGFLLALPAP-LVTEPAEAHRALEAAGVGPALPEPTLPL 213
QY 143 VRSQRFKFKYTKINKTP-LPPRPAP-LTLKKYQPL-----PPAPESASVAPAP 193
DB 214 HNALTR-FLDHLGAVAGRAPAPGAVVALGATFPPLRLRAPPPSPPGG-----APDG 268
QY 194 TPEEVQGRFORSANDSFVLAEEESHQTKRESSCPSSNONTOKSPPAIASSYMPK 253
DB 269 TEE-----TPFPALVETKLQEHLEBQEVAPALPPKPKTKPA----- 308
QY 254 HSIQARDHTGSMOHCAPRCQAASHSPMLPENTNSKPKPTKDEKDWQNEYIGE 313
DB 309 -----PTGLANGSP-----PSLQDA---EWYKCD 330
QY 314 YSRQAVEDVLMKKNKDTFLVRDCSTKSKAEPVLYVFGNKVYKIRFLESNOQFALG 373
DB 331 ISREEVNKKL-RTPDGTFLVRDASSKIQGE-YTLTRKGGN--NKLIVFHRDGHGFS 386
QY 374 TGLRGNEMFDSVEDIIEHYTY 394
DB 387 EPL-----TFCSYVDLTHYRH 403
RESULT 4
P85B_HUMAN
ID P85B_HUMAN STANDARD; PRT; 728 AA.
AC 000459;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
DE p85-beta subunit) (Ptdins-3-kinase p85-beta).
GN PIK3R2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98241181; PubMed=9582025;
RA Janssen J.W.G., Schlichthoff L., Bartram C.R., Schulz A.S.;
RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase
RT p85beta subunit and HMGRF8, a putative deubiquitinating enzyme.";
RL Oncogene 16:1767-1772(1998).
CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: X80907; CAA56868.1; -
DR HSSP: P23727; 2PNB.
DR Genew: HGNC:8980; PIK3R2.
DR MIM: 603157;
DR InterPro: IPR001720; PI3kinase_P85.
DR InterPro: IPR000198; RhogAP.

DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF000017; SH2; 2.
 DR Pfam: PF000018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00678; PI3KINSEP85.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00325; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS0001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 4 80 SH3.
 FT DOMAIN 125 258 RHO-GAP.
 FT DOMAIN 330 425 SH2 1.
 FT DOMAIN 622 716 SH2 2.
 SQ SEQUENCE 728 AA; 81624 MW; 80C2AF244977346B CRC64;

Query Match 6.3%; Score 146; DB 1; Length 728;
 Best Local Similarity 22.9%; Pred. No. 0.027;
 Matches 87; Conservative 48; Mismatches 147; Indels 98; Gaps 18;

OY 46 PLDPHRLNLAGVGGKCNNDYEDPEFO--LLKAMPMSK-----ILPAPPIQE 93
 DB 95 PLPAPRPGAGPEPGTLPLDPEQFSPDVAAPLLVKLVLEAIRTGIDSESHYRPLPAPR 154
 OY 94 SEYA-----DTRYFQDMMEAPLL-LP-----PKASVSTEGQTDVAKMTQLEEDKPTF 140
 DB 155 TDMSLSDVDQMDTALADGKISFLALPAPLVTPAPASAEARALAEAGVPAPALPEPTL 214
 OY 141 KDVSQRKGFYKTKINKTPLEPPR--PAI-TLPKKYQPL--PPAPPESSAYAPAPKT 194
 DB 215 PLHRLTLR-FLDHLGNVARRAPALGPVAVRNLGATFGLLRAPPPSSPPGAPGGS 273
 OY 195 FPEVGRGRORAKDFSKVLGAEEESHQOTPESSCPSSNONTOKSPALASSVMPGKH 254
 DB 274 EP-----SPDFPALVLEKLEQHELEQEVAPALPPKPPAKPA----- 312
 OY 255 SIQARDHTGSMQHCPAQCQAASHSPMLPYENTNSEKPPDTPKDEKDVQNMENYIGEY 314
 DB 313 -----PTVL-----ANGSP-----PSLQDA--EYWGDI 335
 OY 315 SRQAVEDVLMKNDGTFLVRCSTKSKAEPYLVVFGNKVYNYKIRFLESNOQFALGT 374
 DB 336 SREEVNEKL-RDTPDGTFLVRDASSKIQGE-YTILRKGN--NKLIVFHRDGHYGESE 391
 OY 375 GLRGEMEDSVEDITEHTY 394
 DB 392 PL---TFCSVVDLINHYRH 407

RESULT 5
 P85B_RAT STANDARD; PRT; 722 AA.
 AC Q63788;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 DE p85-beta subunit) (PtdIns-3-kinase p85-beta).
 GN PIK3R2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Jistar; TISSUE=Brain;
 RX MEDLINE=96214979; Pubmed=8621382;
 RA Inukal K., Anal M., Vanbrede E., Hosaka T., Katagiri H., Funaki M.,

RA Fukushima Y., Ogihara T., Yazaki Y., Kiyuchi M., Oka Y., Asano T.;
 RT "A novel 55-kDa regulatory subunit for phosphatidylinositol 3-kinase
 RT structurally similar to p55PIK is generated by alternative splicing
 RT of the p85alpha gene.";
 RL J. Biol. Chem. 271:5317-5320(1996).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATE
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D64046; BAA10926.1; -.
 DR HSSP: P23727; 2PMB.
 DR InterPro: IPR001720; PI3KINase_P85.
 DR InterPro: IPR001980; RhogAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00678; PI3KINSEP85.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00325; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 4 80 SH3.
 FT DOMAIN 125 257 RHO-GAP.
 FT DOMAIN 324 419 SH2 1.
 FT DOMAIN 616 710 SH2 2.
 SQ SEQUENCE 722 AA; 81328 MW; 1208368B9F6FC95 CRC64;

Query Match 6.2%; Score 143.5; DB 1; Length 722;
 Best Local Similarity 22.1%; Pred. No. 0.038;
 Matches 80; Conservative 39; Mismatches 120; Indels 123; Gaps 17;

OY 138 PTFKDVRSQR--FKGFKYTKINKTPLEPPRAITLPKKYQPLPAP--PEES----- 185
 DB 58 PGFNEKTRQRQDPEGTVEEFGPVALARPGPR--PGCPRLPAPRDPDGSESGHTLASL 114
 OY 186 SAYPAPRTPPEVGRGRORAKDFSKVLGAEEESHQOTPESSCPSSNONTOK----- 239
 DB 115 AEQSPESPAPPI-----LVKLEIAIEQAEIDSEFYSRPELPAPRTDWSLSDLEQMDR 167
 OY 240 -----SPPAIASSSY-----NPGKHSIQAR---D 260
 DB 168 TTYLDAVKGFLPALPAVVPPEAASEAYRAMREYTGVLGPLEPPLPLQALTLRLFLQ 227
 OY 261 HTGSMQH--CPAQRCAAS--HSPPRL-----PYENTNSEKPPDTPK-----D 300
 DB 228 HLGNVARRAPSPAVVAHALASAFGLPILRAPPPGEGDGESEPAADFVLLERLVQEHVD 287
 OY 301 EKD-----VMQNEWYIGEGSRAVVEDVLMKENDGTG 332
 DB 288 EODTAPALPPPKSKVPAATYALANGSTSLQDAEYWGDISREEVNEKL-RDTPDGTG 346
 OY 333 LVRCSTKSKAEPYLVVFGNKVYNYKIRFLESNOQFALGTGLRGEMEDSVEDITEHTY 392

```

Db      347 LVKDASSKIQGE-YTLTKRKGNN--NKLTKVHHDGHGYSFSEPL-----TFCSVVELISHY 399
QY      393 TY 394
Db      400 RH 401

RESULT 6
FCR_PSVGR STANDARD: PRT: 545 AA.
AC P00544;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tyrosine-protein kinase transforming protein FGR (EC 2.7.1.112).
GN V-FGR OR SRC-2.
OS Feline sarcoma virus (strain Gardner-Rasheed).
OC Viruses; Retroid viruses; Retroviridae; Mammalian type C retroviruses.
OX NCBI_Taxid=11775;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84097512; PubMed=6318314;
RA Naharro G., Robbins K.C., Reddy E.P.;
RT "Gene product of v-tyr onc: hybrid protein containing a portion of
RT actin and a tyrosine-specific protein kinase.";
RL Science 223:63-66(1984).
CC -1 CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
CC tyrosine phosphate.
CC -1 MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-FGR
CC POLYPROTEIN.
CC -1 SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. SRC
CC SUBFAMILY.
CC -1 SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: X00255; CAA25063.1;
DR PIR: A00653; TVMYR.
DR HSR: P00523; 2PTK.
DR InterPro: IPR004001; Actin.
DR InterPro: IPR004000; Actin_like.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001245; Tyr_pkinase.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF00022; actin; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00401; SH2DOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR ProDom: PD000001; Euk_pkinase; 1.
DR ProDom: PD000093; SH2; 1.
DR SMART: SM00268; ACTIN; 1.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS00406; ACTIN; 1.
DR PROSITE: PS00132; ACTIN_ACT_LIKE; 1.
DR PROSITE: PS50001; SH2; 1.
KM Polyprotein; Tyrosine-protein kinase; Transferase; Phosphorylation;
KM ATP-binding; Oncogene; SH2 domain.
FT DOMAIN 23 157 ACTIN.
FT DOMAIN 167 264 SH2.
FT DOMAIN 286 539 PROTEIN KINASE.
FT NP_BIND 292 300 ATP (BY SIMILARITY).

```

```

FT BINDING 314 314 ATP (BY SIMILARITY).
FT ACT_SITE 405 405 BY SIMILARITY.
FT MOD_RES 435 435 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 545 AA; 61516 MW; 7F4742EB7A7CA13E CRC64;

Query Match
Best Local Similarity 23.6%; Pred. No. 0.034; Length 545;
Matches 66; Conservative 30; Mismatches 96; Indels 88; Gaps 13;

QY 165 RPAITLPKKRYOPLPPAPPESSAYF-----APKPTPEVORGRORSA 207
Db 6 RPAYCPRPLPLPLPTAMEEVAALVINDSGMKRAGFADGAPRAVFEISYGRRHQCV 65
QY 208 -----KDSRYLGAEEESHQHTKPESSCPSSNONTQKSPPIASSSYPGKHSTQAR-- 259
Db 66 MWGMGQRD--SYVGDEAOSKR-----GILTLKY-PIEHGIYTWMD 102
QY 260 DHTGSMQHCPCAPORCAAAASHSPMLPYENTNSEKP-DPTKPKDEK----- 302
Db 103 DMEKITHHTFYEDLRAPEHEHYLL-----TEAPLNPKANEKMTQIMFTFNIPSNV 156
QY 303 ----DVMQNEWYIGEXSRQAVE-DVLMKENKDGTEPLVRDCSTKSKAEPYLVV-----FY 352
Db 157 APVDSIOAEEMWYFGKIGRKDAEROLLSPGNARGAFLVRESETTKGA--YSLSRDWEAR 214
QY 353 GKRYVNVKIRPLESNQGFALGGLGNGEMFDSVEDIIEHY 392
Db 215 GDVVKHYKTRKLDTG-----GYITTTAQNFSVQELVQHY 249

RESULT 7
SCK_HUMAN STANDARD: PRT: 540 AA.
AC P98077; Q9NPJ5; O60230;
DT 01-FEB-1996 (Rel. 33, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Protein SCK (Fragment).
GN SCK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98175964; PubMed=9507002;
RA Nakamura T., Muraoka S., Sanokawa R., Mori N.;
RT "N-Shc and Sck, two neuronally expressed Shc adapter homologs. Their
RT differential regional expression in the brain and roles in
RT neurotrophin and Src signaling.";
RL J. Biol. Chem. 273:6960-6967(1998).
RN [2]
RP SEQUENCE OF 72-499 FROM N.A.
RX MEDLINE=95090462; PubMed=7527937;
RA Kavanaugh W.M., Williams L.T.;
RT "An alternative to SH2 domains for binding tyrosine-phosphorylated
RT proteins.";
RL Science 266:1862-1865(1994).
RN [3]
RP SEQUENCE OF 72-328 FROM N.A.
RX Lamerdin J.E., McCreedy P.M., Skowronski E., Viswanathan V.,
RA Burkhardt-Schultz K., Gordon L., Dias J., Brower A., Ramirez M.,
RA Stillwagen S., Phan H., Velasco N., Do L., Regala W., Terry A.,
RA Ganes J., Dangnan L., Erler A., Christensen M., Georgescu A.,
RA Avila J., Liu S., Attix C., Andreise T., Trankiel M.,
RA Amico-Keller G., Coefield J., Duarte S., Lucas S., Bruce R.,
RA Thomas P., Quan G., Krommiller B., Arellano A., Sanders C., Ow D.,
RA Nolan M., Trong S., Kobayashi A., Olsen A.S., Carraro A.V.;
RT "Sequence analysis of a 3.5 kb contig in human 19p13.3 containing a
RT serine protease gene cluster.";
RN Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
RP [4]
SEQUENCE OF 149-540 FROM N.A.

```

RA Carlin L., Estivill X., Sunoy L., Escarceller M.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY IN LIVER, ALSO PRESENT IN BRAIN.
 CC -1- SIMILARITY: CONTAINS 1 PID DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: TO SHC TRANSFORMING PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AB001451; BAA25798.1; -
 DR EMBL: AC006124; AAC97988.1; -
 DR EMBL: AL360254; CAB96175.1; -
 DR HSSP: P29353; IMIL.
 DR MIM: 605217; -
 DR InterPro: IPR000050; PID_domain.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00640; PID; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRINTS: PR00629; SHCPIDOMAIN.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00462; PTB; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: PS01179; PID; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR SH2 domain.
 KW NON_TPR
 FT DOMAIN 1 1 PID.
 FT 105 287 SH2.
 FT 445 536
 FT 72 72 A -> P (IN REF. 2).
 FT CONFLICT 72 72
 SQ SEQUENCE 540 AA; 57623 MW; 0D83700910B1A5DB CRC64;
 Query Match 6.0%; Score 140; DB 1; Length 540;
 Best Local Similarity 22.3%; Pred. No. 0.046;
 Matches 96; Conservative 55; Mismatches 185; Indels 94; Gaps 18;
 QY 13 GFGDLRFQWNLKNSRNPSTLSAKGRCRAVLEPLDPRRLAGVPGGKCNNDY-- 69
 DB 169 GKSLRRAGNSISIHSTIDGLSVPRATROY--ANHHMPSISASGSD--TMDTIVAY 224
 QY 70 --EDPERQ---LKAMPMSKILPARPIQES--EYADTRFGQDMMEAR-LILPKKASYST 120
 DB 225 VAKDPINORACHILECEGL---AQSISTVGOAFELREFKQYLHSPKVALPPELAGP 280
 QY 121 ERQTRDVRMQLQLE-----EVDKPTFKDVSQRKFGKTKYKINKTPLRPPRAPATL 170
 DB 281 EESAMGDEEDSLHNYSINISIPKPEPLGLVDSRLALTPCALALDQGPSPSLRDGSL 340
 QY 171 PKKQPLPAPPESSAVFAKPTFEVQ---RGPRQSAKDFSRVLAEESSHHTQKPE 227
 DB 341 PMVDGSGTAPRGD-----YVQADAKRP-----PDHEHLVYNTQGLDAPEPE 384
 QY 228 SSCSSNONTQKSPPAISSYMEGKHSIQARDHTGSMQHCPAORCAAAASHSPMLPEYE 287
 DB 385 DS-----PKKDLFDWRPFEDALK---LHECSVAAGVTAAPLEPLE 420
 QY 288 NTNSEKPD---PTKPDENVQNMWYIGEYSRAVEDYLMKENNDGTFVLVADCTSKSAE 344
 DB 421 DQWDSPTRRAPVAPTEQQLQOEWPYHGMRRAERKL---RADGGFLVDSVT--NPG 475
 QY 345 PYVLVVFNGKVVYWKIRFLESNOQFALGTGLRCNEMDSVEDIETHYPIILLIDCKD 404
 DB 476 QYVLTLHMAQPKHLLLVDP-----GVVRTKDVLFESISHLIDH-----LQNGQP 522
 QY 405 KAARRKOCYL 414
 DB 523 IVAASESLHL 532

RESULT 8
 ID 3BP2_HUMAN STANDARD; PRT; 561 AA.
 AC P78314; O15373; 000500; P78315;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE SH3 domain-binding protein 2 (3BP-2).
 GN SH3BP2 OR 3BP2 OR RES4-23.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=97446134; PubMed=9299232;
 RA Bell S.M., Shaw M., Jou Y.-S., Myers R.M., Knowles M.A.;
 RT "Identification and characterization of the human homologue of
 RT SH3BP2, an SH3 binding domain protein within a common region of
 RT deletion at 4p16.3 involved in bladder cancer.";
 RL Genomics 44:163-170(1997).
 RN [2]
 RP MEDLINE=98403881; PubMed=9734812;
 RA Hadano S., Ishida Y., Ikeda J.-E.;
 RT "The primary structure and genomic organization of five novel
 RT transcripts located close to the Huntington's disease gene on human
 RT chromosome 4p16.3.";
 RL DNA Res. 5:177-186(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (LONG ISOFORM), AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE=98403881; PubMed=9734812;
 RA Hadano S., Ishida Y., Ikeda J.-E.;
 RT "The primary structure and genomic organization of five novel
 RT transcripts located close to the Huntington's disease gene on human
 RT chromosome 4p16.3.";
 RL DNA Res. 5:177-186(1998).
 RN [4]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RC TISSUE=tonsil;
 RA Gokemeijer J., Deligiannidis K.E., Ligris K., Ernst T.J.;
 RT "3BP2 binds to phosphatidylinositols: linking the hemopoietic tyrosine
 RT kinase c-Fes to the cytoplasmic membrane in a phosphorylation
 RT dependent mechanism.";
 RL Blood 88:473A-473A(1996).
 RN [5]
 RP VARIANTS CRBM GLN-415; PRO-415; ARG-418; HIS-418; LEU-418; ARG-420 AND
 RP GLU-420.
 RX MEDLINE=21275962; PubMed=11381256;
 RA Ueki Y., Tizabi V., Santanna C., Fukai N., Maulik C., Garfinkle J.,
 RA Nimniya C., dokamaral C., Peters H., Hbal M., Rhee-Morris L.,
 RA Doss J.B., Kreisberg S., Olsen B.R., Reichemberger E.;
 RT "Mutations in the gene encoding c-Abl-binding protein SH3BP2 cause
 RT cherubism.";
 RL Nat. Genet. 28:125-126(2001).
 CC -1- FUNCTION: BINDS DIFFERENTIALLY TO THE SH3 DOMAINS OF CERTAIN
 CC PROTEINS OF SIGNAL TRANSDUCTION PATHWAYS. BINDS TO
 CC PHOSPHATIDYLINOSITOLS; LINKING THE HEMOPOIETIC TYROSINE KINASE
 CC C-FES TO THE CYTOPLASMIC MEMBRANE IN A PHOSPHORYLATION DEPENDENT
 CC MECHANISM.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms: a long form (shown here) and a
 CC short form: are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in a variety of tissues including
 CC lung, liver, skeletal muscle, kidney and pancreas.
 CC -1- DISEASE: Defects in SH3BP2 are the cause of cherubism (CRBM), an
 CC autosomal dominant inherited syndrome. It is characterized by
 CC excessive bone degradation of the upper and lower jaws, which
 CC often begins around three years of age. It is followed by
 CC development of fibrous tissue masses, which causes a
 CC characteristic facial swelling.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

Db 429 EKPPOPSADRGCDSDDEYEVPLPNSVFNVTTESECEVERLFATSPRGEPDGLYCIR 488
 336 DCSSTKAEPLYLVAF--YGNKVVYVVKIRFLIESNOQFALGTLRGNEFSDVEDIIIEY 392
 489 NSTSTSGK---VLVYWDETSNKVRNRYRI--PEKDSKFY----LEGEVLFSVSGSNVEHY 538

RESULT 9
 BMX_HUMAN
 ID BMX_HUMAN STANDARD: PRT: 675 AA.
 AC P51813; 012871; 060564;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cytoplasmic tyrosine-protein kinase BMX (EC 2.7.1.112) (Bone marrow kinase BMX) (Epithelial and endothelial tyrosine kinase) (ETK) (NTRK38).
 DE BMX.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC TISSUE=BONE MARROW;
 RX MEDLINE=95060827; PubMed=7970727;
 RA Tmagnone L., Lahtinen I., Mustonen T., Vitaneva K., Francis F., Muscatelli F., Allitalo R.P., Smith C.T., Larsson C., Allitalo K.;
 RT "BMX, a novel nonreceptor tyrosine kinase gene of the BTK/TK/TEC/TXK family located in chromosome xp22.2.";
 RL Oncogene 9:3683-3688(1994).
 RN (2)
 RP SEQUENCE FROM N.A., FUNCTION, AND INDUCTION.
 RC TISSUE=Prostate;
 RX MEDLINE=9618236; PubMed=9520419;
 RA Qiu Y., Robinson D., Prellow T., Kung H.-J.;
 RT "Etk/Bmx, a tyrosine kinase with a pleckstrin-homology domain, is an effector of phosphatidylinositol 3'-kinase and is involved in interleukin 6-induced neuroendocrine differentiation of prostate cancer cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:3644-3649(1998).
 RN (3)
 RP SEQUENCE OF 536-599 FROM N.A.
 RC TISSUE=Blood;
 RX Fuortes M.;
 RL Thesis (1994), Cornell University, U.S.A.
 CC -1- FUNCTION: ACTIVITY IS REQUIRED FOR INTERLEUKIN 6 (IL-6) INDUCED DIFFERENTIATION. MAY PLAY A ROLE IN THE GROWTH AND DIFFERENTIATION OF HEMATOPOIETIC CELLS. MAY BE INVOLVED IN SIGNAL TRANSDUCTION IN ENDOCARDIAL AND ARTERIAL ENDOTHELIAL CELLS.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine = ADP + protein tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
 CC -1- TISSUE SPECIFICITY: PREFERENTIALLY EXPRESSED IN EPITHELIAL AND ENDOTHELIAL CELLS.
 CC -1- INDUCTION: ACTIVATED BY IL-6 THROUGH PHOSPHATIDYLINOSITOL 3-KINASE (PI3-KINASE) PATHWAY. IT IS LIKELY THAT ACTIVATION OCCURS THROUGH BINDING OF PHOSPHOSITIDES TO THE PH DOMAIN.
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

DR EMBL: X83107; CAAS8169.1; -
DR EMBL: U08341; AAA17744.1; -
DR EMBL: AF045459; AAC08966.1; ALT_INIT.
DR HSSP: Q06187; 1B55.
DR Genew: HGNC:1079; BMX.
DR MIM: 300101; -
DR InterPro: IPR001562; BTK.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001849; PH.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF00069; pkinase; 1.
DR Pfam: PF00169; PH; 1.
DR Pfam: PF00729; BTK; 1.
DR PRINTS: PR00402; TECBTKDOMAIN.
DR PRINTS: PR00109; TYRKINASE.
DR PRODOM: PD000001; Euk_pkinase; 1.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00107; BTK; 1.
DR SMART: SM00233; PH; 1.
DR SMART: SM00232; SH2; 1.
DR SMART: SM00219; TYRK; 1.
DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50003; PH_DOMAIN; 1.
DR Transferrase: Tyrosine-protein kinase; Phosphorylation;
KW ATP-binding; SH3 domain; SH2 domain.
FT DOMAIN 4 111 PH.
FT DOMAIN 196 289 SH3.
FT DOMAIN 296 392 SH2.
FT DOMAIN 417 675 ATP (BY SIMILARITY).
FT N-BIND 423 431 ATP (BY SIMILARITY).
FT BINDING 445 445 ATP (BY SIMILARITY).
FT ACT_SITE 536 536 ATP (BY SIMILARITY).
FT MOD_RES 566 566 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT DOMAIN 251 255 POLY-SER.
FT DOMAIN 282 285 POLY-SER.
FT DOMAIN 286 289 POLY-GLU.
FT CONFLICT 597 597 A -> S (IN REF. 3).
SQ SEQUENCE 675 AA; 78010 MW; CB2382A3BD02599 CRC64;

Query Match 5.7%; Score 133; DB 1; Length 675;
Best Local Similarity 24.08; Pred. No. 0.16; Indels 48; Gaps 11;
Matches 62; Conservative 41; Mismatches 107;

QY 166 PAILPKKYOPLPPAPPESSAYEPKPTPE-VQGRPR-----QSAKDFSVLGA--- 216
DB 141 PGCTLMAYANLHRAVNEKHRV---PTFPDRLKIPRAVPVLKMDAPSSSTTLAQYDN 196
QY 217 EEEHHOTKPESSCPSSNQNTQKSPAIASS-----SYMGC-----KHSIQAR 259
DB 197 ESKKNYSSQPPSSSTSLAQYDSNKKIYGSQPNFMQYIPREDDPPDMWQVAKLLSSSSSE 256
QY 260 DHTSMOHCPRQCAASHPRMLPYENTNSEKPDPTKPEKDWQNGWNTIGESQAV 319
DB 257 DVASSNKK-----ERNVNHSTKISWEPSSSE-----EEULDDYDFAGIISQS 306
QY 320 EDVLAKENKDGTLVRCSTKSKAPVYLVEF---YGNKYVNVKIRPLESNOQFALGTGL 376
DB 307 EQLLRQKGEKAFVNR---SSQVGMVTVLSFSAVNDKGTAKHVVHTVAEKKL--YL 361
QY 377 RGNEMFDSVEDIIEHTY 394
DB 362 AENYCFDSIPRLHYHQH 379

```

```

DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
DE p85-alpha subunit) (Pcdins-3-kinase p85-alpha) (PI3K).
GN PIK3R1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS P85-ALPHA AND P55-ALPHA).
RC STRAIN=Histar; TISSUE=Brain;
RX MEDLINE=96214979; PubMed=8621382;
RA Inukai K., Anal M., van Breda E., Hosaka T., Katagiri H., Funaki M.,
RA Fukushima Y., Ogihara T., Yazaki Y., Kikuchi M., Oka Y., Asano T.,
RT "A novel 55-kDa regulatory subunit for phosphatidylinositol 3-kinase
RT structurally similar to p55PIK is generated by alternative splicing
RT of the p85alpha gene."
RL J. Biol. Chem. 271:5317-5320(1996).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM P50-ALPHA).
RC TISSUE=Liver;
RX MEDLINE=97218222; PubMed=9065454;
RA Inukai K., Funaki M., Ogihara T., Katagiri H., Kanda A., Anal M.,
RA Fukushima Y., Hosaka T., Suzuki M., Shin B., Taketa K., Yazaki Y.,
RA Kikuchi M., Oka Y., Asano T.,
RT "p85alpha gene generates three isoforms of regulatory subunit for
RT phosphatidylinositol 3-kinase (PI 3-kinase), p50alpha, p55alpha, and
RT p55alpha, with different PI 3-kinase activity elevating responses to
RT insulin."
RL J. Biol. Chem. 272:7873-7882(1997).
RN [3]
RP SEQUENCE FROM N.A. (ISOFORM P50-ALPHA).
RC TISSUE=Liver;
RX MEDLINE=97079666; PubMed=8921377;
RA Furman D.A., Cantley L.C., Carpenter C.L.;
RT "Structural organization and alternative splicing of the murine
RT phosphoinositide 3-kinase p85 alpha gene."
RL Genomics 37:113-121(1996).
CC - FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
CC - SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC - ALTERNATIVE PRODUCTS: 3 ISOFORMS: P85-ALPHA (SHOWN HERE), P55-
CC ALPHA AND P50-ALPHA; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC - TISSUE SPECIFICITY: THE P85-ALPHA ISOFORM IS WIDELY EXPRESSED.
CC EXPRESSION OF THE P55-ALPHA ISOFORM IS HIGHEST IN BRAIN AND
CC SKELETAL MUSCLE. THE P50-ALPHA ISOFORM IS ABUNDANT IN LIVER WITH
CC LOWER LEVELS IN BRAIN AND MUSCLE.
CC - SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC - SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC - SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D64045; BAA18932.1; -
DR EMBL: D64048; BAA18933.1; -
DR EMBL: U50412; AAC52846.1; -
DR EMBL: D78486; BAA24426.1; -
DR HSSP: P23727; 1BFI.
DR InterPro: IPR001720; PI3kinase_P85.
DR InterPro: IPR000198; RhogAP.
DR InterPro: IPR000980; SH2.

```

DR InterPro; IPR001452; SH3.
 DR Pfam; PF00017; SH2; 2.
 DR Pfam; PF00018; SH3; 1.
 DR Pfam; PF00620; RHO GAP; 1.
 DR PRINTS; PR00678; PTKINASEP5.
 DR PRINTS; PR00401; SH2DOMAIN.
 DR ProDom; PD000066; SH3; 1.
 DR ProDom; PD000093; SH2; 2.
 DR SMART; SM00324; RHO GAP; 1.
 DR SMART; SM00252; SH2; 2.
 DR SMART; SM00326; SH3; 1.
 DR PROSITE; PS50001; SH2; 2.
 DR PROSITE; PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat; Alternative splicing.
 KW DOMAIN 3 79 SH3.
 FT DOMAIN 333 428 SH2 1.
 FT DOMAIN 624 718 SH2 2.
 FT VARSPLIC 1 270 MISSING (IN ISOFORM P55-ALPHA).
 FT VARSPLIC 271 304 VLFRPAAASDNTMHLKAVELLISAEWSEKOPA -> MYT
 FT VARSPLIC 271 304 TVWTMEDLDIECAKTDINCCTDIMEFYEMDP (IN
 FT VARSPLIC 1 300 ISOFORM P55-ALPHA).
 FT VARSPLIC 301 306 MISSING (IN ISOFORM P50-ALPHA).
 SQ SEQUENCE 724 AA; 83531 MW; 95C65CF612873B84 CRC64.

Query Match 5.7%; Score 133; DB 1; Length 724;
 Best Local Similarity 22.2%; Pred. No. 0.18;
 Matches 88; Conservative 46; Mismatches 155; Indels 108; Gaps 19;

QY 38 GRCRAVLEPLPDHR--RNLAGVPGGCKNSNDYEDPEFOLKAWPSMKILPARPIOESE 95
 DB GRKR-ISPPTKPRPRPRPLFVAPGSSKTEA--DTEQPVLT-----PDIAEQ 121
 QY 96 YATRFQDMMKAPLLPRASVS-----TERQDVMQLMQLVEVDKPTF----- 140
 DB 122 FAPPD-----VAPPLKLLLEALEKGLCECTIKTQSSNPALROLDCDDPSVDLV 176
 QY 141 --KDVRSOREKFGKYTKINTPLPPPPATLTKKYQPL-PPAPPESSAYFAK----P 193
 DB 177 FDEHVLADAKRY-----LADLPNPIYIPVAVYNEMMSLAQVPPSSDYIQLKLIRSP 230
 QY 194 TFEVORGPRQSAKDFRYLGAEBESHQTKPESSC-----PSSNQNTQKSPAT 244
 DB 231 NIHQYWLTLQYLLKHFKFLSOASSKLLNARALSEIFSHVLPFRPAASDNT----- 284
 QY 245 ASSSYMPGKHSIOARDHTGSMQHPQRCQAASHSPRMPIYENTNEKRDPKP----- 299
 DB 285 -----HLKAVELLISAEWSEKOPAPA-----LP-----KPKPKTISIAN 319
 QY 300 ----DEKDVQNMENYIGEYSRQAVEDVLMKENKDGTEFLVRCSTSKSAEPYLVAVFYGNK 355
 DB 320 NSMNNMSIQDAEMYMWDISREYNEKL-RDTADGTFVLVADASTKMGD-YTLTLRKGCGN 377
 QY 356 VYVVKIRFLESNOQFALGTGLRGNEFDSVEDIIEHY 392
 DB 378 --NKLIRKIFHRDGKYGSDPL-----TENSVELINHY 408

RESULT 11

GRAP_HUMAN STANDARD; PRT; 217 AA.
 AC 013588;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE GRB2-related adaptor protein.
 GN GRAP.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=96218119; PubMed=8647802;
 RA Feng G.-S., Ouyang Y.-B., Hu D.-P., Shi Z.-Q., Genz R., Ni J.;
 RT "Grap is a novel SH3-SH2-SH3 adaptor protein that couples tyrosine
 RT kinases to the Ras pathway.";
 CC J. Biol. Chem. 271:12129-12132(1996).
 CC -1- FUNCTION: COUPLE SIGNALS FROM RECEPTOR AND CYTOPLASMIC TYROSINE
 CC KINASES TO THE RAS SIGNALING PATHWAY.
 CC -1- SUBUNIT: ASSOCIATES THROUGH ITS SH2 DOMAIN WITH LIGAND-ACTIVATED
 CC RECEPTORS FOR STEM CELL FACTOR (KIT) AND ERYTHROPOIETIN (EPOR).
 CC ALSO FORMS A STABLE COMPLEX WITH THE RAS GUANINE NUCLEOTIDE EXCHANGE FACTOR SOS1,
 CC PRIMARILY THROUGH ITS N-TERMINAL SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC -1- SIMILARITY: BELONGS TO THE GRB2 / SEM-5 / DRK FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; U52518; AAC50541.1; -
 DR HSSP; 060631; 1GBQ.
 DR Genew; HGNC:4562; GRAP.
 DR MTM; 604330; -
 DR InterPro; IPR000980; SH2.
 DR InterPro; IPR001452; SH3.
 DR Pfam; PF00017; SH2; 1.
 DR Pfam; PF00018; SH3; 2.
 DR PRINTS; PR00452; SH3DOMAIN.
 DR ProDom; PD000066; SH3; 2.
 DR ProDom; PD000093; SH2; 1.
 DR SMART; SM00252; SH2; 1.
 DR SMART; SM00326; SH3; 2.
 DR PROSITE; PS50001; SH2; 1.
 DR PROSITE; PS50002; SH3; 2.
 KW SH2 domain; SH3 domain; Repeat.
 FT DOMAIN 1 58 SH3 1.
 FT DOMAIN 60 152 SH2.
 FT DOMAIN 156 215 SH3 2.
 SQ SEQUENCE 217 AA; 25336 MW; 09FEC2F3BAC0FAF8 CRC64;

Query Match 5.7%; Score 132.5; DB 1; Length 217;
 Best Local Similarity 27.7%; Pred. No. 0.048;
 Matches 36; Conservative 24; Mismatches 41; Indels 29; Gaps 4;

QY 293 KPDPKPDQDMQNMENYIGEYSRQAVEDVLMKENKDGTEFLVRCSTSKSAEPYLVAVFY 352
 DB 56 KHPH-----WYSGRISQLAEELIKMKNHGLAIRL--SESSGPEFSVYNY 101
 QY 353 GNKYVYVVKIRFLESNOQFALGTGLRGNEFDSVEDIIEHYTYFPILLIDGDKAARRKQC 412
 DB 102 GDQVQHFVLRKASGKVFEL-----WEKFNLSLNELVDFRTTYI-----AKKRQI 146

QY 413 YLTQPLPLAR 422

DB 147 FLRDEPLLR 156

RESULT 12

CRK_DROME STANDARD; PRT; 271 AA.
 ID CRK_DROME
 AC 09XYM0;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Adapter molecule ctk.
 GN CRK OR CG1387.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;

CC Insecta: Pterygota; Neoptera; Endopterygota: Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila;
XX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Embryo;
RX MEDLINE=99173888; PubMed=10072777;
RA Galletta B.J., Niu X.-P., Erickson M.R., Admayr S.M.;
RT "Identification of a Drosophila homologue to vertebrate Crk by
RL interaction with MBC.";
RN Gene 228:243-252(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt J.C., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baltow D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Bencs P.V., Bernan B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Boutek J., Brockstein P., Brotlier P.,
RA Butkus K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dublin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwan C.,
RA Jalali M., Kallush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C., Kraft C., Kravitz S., Kulp D., Lal Z.,
RA Laoko P., Lecl Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mottel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi D.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Schelker F., Shen H.,
RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.);
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: ADAPTER PROTEIN WHICH INTERACTS WITH C-TERMINAL PORTION
OF MBC, HOMOLOG OF HUMAN DOCK180. MAY PLAY A ROLE IN CELLULAR
PROCESSES THROUGHOUT DEVELOPMENT.
CC -1- TISSUE SPECIFICITY: EMBRYONIC ZYGOTIC EXPRESSION IS SEEN IN
INVAGINATING PRESUMPTIVE MESODERM AND ECTODERMALLY DERIVED TISSUES
DURING GASTRULATION. AT STAGE 8, EXPRESSION IS ALSO SEEN IN
ANTERIOR AND POSTERIOR MIDGUT AND CEPHALIC FURROW. BY STAGE 9,
EXPRESSON IS HIGHEST IN VISCERAL MESODERM OF ANTERIOR AND
POSTERIOR MIDGUT, VENTRAL NERVE CORD AND SOMATIC MESODERM.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY
THROUGHOUT EMBRYOGENESIS, DECLINES DURING LARVAL STAGES AND
REAPPEARS DURING PUPATION.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

[illegible]

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X51780; CAA36075.1; -
DR EMBL: X59720; CAA42253.1; -
DR PIR: S19503; LBY.
DR HSSP: P29355; ISEM.
DR SGD: S0000684; ABP1.
DR InterPro: IPR002108; Actbind-cofin.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00018; SH3; 1.
DR Pfam: PF00241; cofilin.ADF; 1.
DR PRINTS: PR00452; SH3DOMAIN.
DR ProDom: PD000066; SH3; 1.
DR SMART: SM00102; ADF; 1.
DR SMART: SM00326; SH3; 1.
DR PROSITE: PS50002; SH3; 1.
DR Cytoskeleton; Actin-binding; SH3 domain; Repeat.
KW SIMILAR 1 142 TO YEAST COFILIN.
FT DOMAIN 86 97 ACTIN-BINDING (POTENTIAL).
FT DOMAIN 532 592 SH3.
FT DOMAIN 200 575 3 x 10 AA APPROXIMATE REPEATS.
FT REPEAT 200 209 1.
FT REPEAT 436 445 2.
FT REPEAT 566 575 3.
FT REPEAT 575 585 L -> S (IN REF. 1).
FT CONFLICT 312 312 K -> I (IN REF. 1).
FT CONFLICT 312 312 L -> S (IN REF. 1).
SQ SEQUENCE 592 AA; 65576 MW; 39523510704D94AA CRC64;

Query Match 5.7%; Score 132.5; DB 1; Length 592;
Best Local Similarity 21.4%; Pred. No. 0.15; Mismatches 99; Indels 55; Gaps 11;
Matches 54; Conservative 44;

QY 117 SYSTERQTDVMTQLEVDKPTFKVRSQREKGFYTKINKTPL-----PPPR---PAT 168
DB 329 NMTKAEAKRPEVDEDEGEEDPVKDLKS-KFEGLAASEKEEEMENKFAFPPKSSPTL 387
QY 169 TLPRKY-QPLPAPRPRESSYFAPKPTFPEVQGPQRQSKDSRYLGAEEESHQTKPE 227
DB 388 ISKRPKSKQEPKAEAE--QPKTDYKIG-----NPLPGMHIEDNEEPE 433
QY 228 SS-----CPSSN-----ONTQKSPPAIASSSYMPGKHSIQARDHTGSM 265
DB 434 ENDDDDDEDAAPPLPSRNVAASGAPVQKEPEDEIAPS--LPSRNIPAPKQEAAP 491
QY 266 QHCPCARCOAAASHSPRLPYENTNSEKDPPTK-PDEKDVQWQWYIGEYSRQAVEDYLM 324
DB 492 EQAPEEIEEAEAEAPQLPSRSSAAPPPRRATPEKKRKNPMATAEYDYAAD--- 548
QY 325 KENKDGTFVLRD 336
DB 549 ---NELTFEEND 557

RESULT 14
PSC_DROME STANDARD; PRT; 1603 AA.
ID PSC_DROME
AC P35820;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Posterior sex combs protein.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92018190; PubMed=1833647;

RA Brunk B.P., Martin E.C., Sharp E., Adler P.N.;
RT "Drosophila genes Posterior Sex Combs and Suppressor two of zeste
RT encode proteins with homology to the murine bml-1 oncogene";
RL Nature 353:351-353(1991).
CC -! FUNCTION: THE POLYCOMB GROUP (PC-G) GENES ARE NEEDED TO MAINTAIN
CC EXPRESSION PATTERNS OF THE HOMEOBOX SELECTOR GENES OF THE
CC ANTENNAPEDIA (ANTP-C) AND BITHORAX (BX-C) COMPLEXES, AND HENCE FOR
CC THE MAINTENANCE OF SEGMENTAL DETERMINATION.
CC -! SUBCELLULAR LOCATION: Nuclear (Probable).
CC -! SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: X59275; CAA1965.1; -
DR PIR: S17983; S17983.
DR FlyBase: FBgn0005624; Psc.
DR InterPro: IPR001841; Znf_ring.
DR Pfam: PF00097; zf-C3HC4; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00518; ZF_RING_1; 1.
DR PROSITE: PS50089; ZF_RING_2; 1.
KW Zinc-finger; Developmental protein; DNA-binding; Nuclear protein.
FT DOMAIN 47 53 POLY-THR.
FT DOMAIN 83 88 POLY-THR.
FT DOMAIN 91 98 POLY-THR.
FT DOMAIN 145 152 POLY-THR.
FT DOMAIN 184 202 POLY-SER.
FT ZN_RING 265 304 RING-TYPE.
FT DOMAIN 642 651 POLY-SER.
FT DOMAIN 1066 1069 POLY-GLY.
FT DOMAIN 1185 1189 POLY-PRO.
FT DOMAIN 1214 1217 POLY-PRO.
FT DOMAIN 1391 1396 POLY-PRO.
FT DOMAIN 1458 1461 POLY-ALA.
FT DOMAIN 1517 1520 POLY-GLY.
SQ SEQUENCE 1603 AA; 169999 MW; 77024E409736473 CRC64;

Query Match 5.7%; Score 132; DB 1; Length 1603;
Best Local Similarity 21.2%; Pred. No. 0.52; Mismatches 148; Indels 98; Gaps 18;
Matches 80; Conservative 51;

QY 26 KNSWPSLSSAKRCR-----AVLEPLDHRRLAGVPG-----GECNSNDYEDPEFOL 76
DB 641 KSSSSSSSSSGERKRKSPPLVPLTIRTERIMSPSGVSTLSPVTSAGAFSDPKSEF 700
QY 77 LKAMP-----SKILPAR-IQSEVADR--YFQD--MNEAPLLPPK 115
DB 701 LKSFALKPIKVKVSPERTLNRAITPPSPSVOASPKSGNNLDLSILMKPSCMPK 760
QY 116 ASYSTERQTD---VMTQLEVDKPT--FK---DVSORFQKFKYTKINKTPLPP-R 165
DB 761 SIASSKRKSKPEYKAVASKKRLSPPLTYDPFKIRLPYTNNGSSSTASPKLEKPLMPPAK 820
QY 166 PAITLPKKYQPL-----PPAPPESSA--YEAP-----KPTPEVQR-GPRQSA 207
DB 821 PPMILAPRKIQPSAQFAPRPPPIHHAGVMSAPGNRTPIAKRYQPLPKASRPFPANIP 880
QY 208 KDFSRVL-----GAEEESHQTKPESSCSSNQWQKSPPAIASSSYMGKH 254
DB 881 NDVNRLLKDACTEIKSIGGSVENSNSAOKPHLYGPKG--ETKMGPPALPATTPSGCNK 938
QY 255 SIQARDITGSMOHC-----AQRCAASHSPMLPYENTNSEKDPPTK 298
DB 939 NVGQAGNLPMASAPPNKNGSSNNLTNLALFNSNMCKCKEAPPGCGRTPMYTPNSIYSPSS 998
QY 299 PDEKDVQWQWYIGEYS 315

DB 999 PQ-----YVPSYN 1006

RESULT 15

P85B_MOUSE STANDARD: PRT: 722 AA.

AC 008908;

DT 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase p85-beta subunit) (Ptdins-3-kinase p85-beta).

GN PI3KR2.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI_Taxid=10090;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=NIH Swiss;

RX MEDLINE=98241181; PubMed=9582025;

RA Janssen J W G, Schleithoff L., Bartram C.R., Schulz A.S.;

RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta subunit and HUMOR8, a putative deubiquitinating enzyme.";

RL Oncogene 16:1767-1772(1998).

CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE KINASERS, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.

CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY) SUBUNIT.

CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.

CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: Y13569; CAI73903.1; ..

DR HSSP: P23727; ZPMB.

DR MGD: MGI:1098772; PI3K2.

DR InterPro: IPR001720; PI3kinase_P85.

DR InterPro: IPR000198; RHO-GAP.

DR InterPro: IPR000980; SH2.

DR InterPro: IPR001452; SH3.

DR Pfam: PF00017; SH2; 2.

DR Pfam: PF00018; SH3; 1.

DR Pfam: PF00620; RHO-GAP; 1.

DR PRINTS: PRO0678; PI3KINASEP85.

DR PRINTS: PRO0401; SH2DOMAIN.

DR PRODOM: PD000066; SH3; 1.

DR PRODOM: PD000093; SH2; 2.

DR SMART: SM00324; RHO-GAP; 1.

DR SMART: SM00252; SH2; 2.

DR SMART: SM00326; SH3; 1.

DR PROSITE: PS50001; SH2; 2.

DR PROSITE: PS50002; SH3; 1.

DR SH3 domain: SH2 domain; Repeat.

FT DOMAIN 4 80 SH3.

FT DOMAIN 125 257 RHO-GAP.

FT DOMAIN 324 419 SH2 1.

FT DOMAIN 616 710 SH2 2.

SO SEQUENCE 722 AA; AB4D49AF30CEC567 CRC64;

Query Match 5.6%; Score 130.5; DB 1; Length 722;

Best Local Similarity 21.5%; Pred. No. 0.26; Mismatches 88; Conservative 44; Mismatches 149; Indels 129; Gaps 20;

OY 53 NLAGYVG-GEKNSNNDYDP--EF-----QLKAMP---SMILPAPRIQSEYADTRF 102

DB 53 NVGMMGFENERGRGDEFGTYVEFLGPVALAPGPRGRPLPARPDGSSSEG-HIL 111

OY 103 QDMME---APLLLP-----KASVTERQTR-----DVRMTOLEEVKPKPFK 141

DB 112 PLAEQFSPPDPAPPLVYKLVLEAIEQAELEDSICYKPELPATRTDMSLSDELQMDRTALY 171

OY 142 DVRSQRFKGF-----KYTKINKT-----PLPPRPATITLPKKYQ 175

DB 172 DA-----VKGFLALPAAVVTPPAALAEVAFALREVAGPVGLVEPPLPLHQLATLRFLLQ 227

OY 176 PL---PPAPPESSAYFAKPTF-----PEVGRGRQRAKQESVYLGAEESHQGT 224

DB 228 HIGVARRAPSPDTVHALASAFGLLRIPSGSGEGDSEVPDFPVYLLERLYQEHVE 287

OY 225 KPSSCPSSNMTOKSPPAIASSSYMPCNHSIOARDHTGSMQHCPAQCQAASHSPML 284

DB 288 EQDAAPPLPAPKPSAKAPR----- 306

OY 285 PYENTNSEKPDPTKPDKDWONENYIGEYSKQAVEDYLMKENKGTPLVRDCSTKSKAE 344

DB 307 PTALANGSP---PSLQDA---EMWGDISRHEVNERL-RDTPDGTFLVRDASSKIOGE 358

OY 345 PYLVVVFYGNKYVNYKIRPLESNOQFALGTGRGNEMFDSVDIIEHTY 394

DB 359 -YTLIRKGN--NLIIVFHRDGHGSEPL-----TCSYVELLSHYRH 401

Search completed: April 21, 2003, 12:36:51

Job time : 13.6547 secs

GenCore version 5.1.4.D5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:14:41 ; Search time 50.706 Seconds

(without alignments)
5933.817 Million cell updates/sec

Title: US-09-856-061-3

Perfect score: 2020
Sequence: 1 ttccagacttcagctgcgc.....catgcacacatacaagaatt 1129

Scoring table:

BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 1816940

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 100%
Listing first 45 summaries

Command line parameters: -DEV=xlp
-MODE=frame_n2p.model
-Q=/cgn2.1/USPTO-spool/US09856061/runat.21042003_113018.3728/app_query.fasta.1.3150
-DB=A.Geneseq.101002 -OFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsnm62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09856061.ecgn1.1.209 @runat.21042003_113018.3728 -NCPU=6 -ICPU=3
-NO_XLPRX -NO_MMAP -LARGEOUTER -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.Geneseq.101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2020	100.0	376	22	AA81071	Human mast cell -sp
2	2009	99.5	428	23	AAE22609	Human MIST splice
3	2009	99.5	443	23	AAE22608	Human MIST protein
4	1694	83.9	353	23	AAE22610	Human MIST splice
5	1132	56.0	435	22	AA81070	Murine mast cell-s
6	349	17.3	84	21	AA603152	Human secreted pro
7	290.5	14.4	449	19	AAW59866	Amino acid sequenc
8	290.5	14.4	456	19	AAW59865	Human protein sequ
9	287.5	14.2	456	19	AAW59865	Amino acid sequenc
10	273	13.5	457	19	AAW59867	Amino acid sequenc
11	238	11.8	46	22	AB8147	Peptide #5653 enco
12	238	11.8	46	22	AB82354	Protein #5353 enco
13	238	11.8	46	22	AAW58783	Human brain expres
14	238	11.8	46	22	AAW71296	Human bone marrow
15	238	11.8	46	22	AAW18973	Peptide #5407 enco
16	238	11.8	46	22	AAW31575	Peptide #5612 enco
17	238	11.8	46	23	ABG41096	Human peptide enco
18	236	11.7	95	20	AAW70587	Human SLP-76, SH2 d
19	214	10.6	40	22	AB83614	Peptide #7120 enco
20	214	10.6	40	22	AB839615	Peptide #7121 enco
21	214	10.6	40	22	AAW60320	Human brain expres
22	214	10.6	40	22	AAW60321	Human brain expres
23	214	10.6	40	22	AAW72853	Human bone marrow
24	214	10.6	40	22	AAW72854	Human bone marrow
25	214	10.6	40	22	AAW33179	Peptide #7216 enco
26	214	10.6	40	22	AAW33180	Peptide #7217 enco
27	214	10.6	40	23	ABG42791	Human peptide enco
28	214	10.6	40	23	ABG42792	Human peptide enco
29	170	8.4	503	22	ABG22854	Novel human diago
30	167	8.3	945	22	AB866077	Drosophila melanog
31	166	8.2	30	22	AB838148	Peptide #5654 enco
32	166	8.2	30	22	AB83355	Protein #5354 enco
33	166	8.2	30	22	AAW58784	Human brain expres
34	166	8.2	30	22	AAW71297	Human bone marrow
35	166	8.2	30	22	AAW18974	Peptide #5408 enco
36	166	8.2	30	22	AAW31576	Peptide #5613 enco
37	166	8.2	30	23	ABG41097	Human peptide enco
38	155	7.7	724	16	AAW85784	Human GRB-1. Homo
39	151.5	7.5	759	22	ABG21983	Novel human diago
40	150.5	7.5	724	13	AAW26060	Growth Factor Rece
41	150.5	7.5	724	21	AAW97143	Human PI3K p85 pol
42	150.5	7.5	724	22	AAW67442	Amino acid sequenc
43	150.5	7.5	724	22	AAW67621	Amino acid sequenc
44	150.5	7.5	724	23	ABG60298	Human lymphoma ass
45	150.5	7.5	968	19	AAW85011	p85alpha-green flo

ALIGNMENTS

RESULT 1
ID AAB81071 standard; Protein: 376 AA.

AC AAB81071:
XX 25-JUN-2001 (first entry)
DE Human mast cell-specific immunoreceptor signal transducer.

KW Mast cell: signal transduction; human; allergic disease; MIST;
KW mast cell-specific immunoreceptor signal transducer.

OS Homo sapiens.

PN JP3146204-B1.

XX 12-MAR-2001.
XX

XX Claim 11, Fig 5, 171pp: English.
 PS
 CC The present invention relates to novel mast cell immunoreceptor signal
 CC transducer (MIST) proteins and polynucleotides encoding such proteins.
 CC MIST sequences of the invention are useful for preventing, treating or
 CC ameliorating a medical condition in mammalian subject. They are useful
 CC for treating an immune disorder involving hyperactivity of B- or T-
 CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
 CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
 CC as targets for therapeutic intervention in immune cell disorders and
 CC inflammatory indications, for diagnosis and/or screening of disorders
 CC or diseases associated with expression of MIST, for screening for
 CC antagonists or inhibitors of the interaction of MIST with cellular
 CC signalling components. They are used in assays that detect activation
 CC or induction of various B and T-cell-related neoplasms or cancers.
 CC Sequences of the invention are also used in gene therapy. The present
 CC sequence is human MIST splice variant protein from clone #7.
 XX

Sequence 428 AA:

Alignment Scores:
 Pred. No.: 5.42e-195 Length: 428
 Score: 2009.00 Matches: 374
 Percent Similarity: 99.73% Conservative: 1
 Best Local Similarity: 99.47% Mismatches: 1
 Query Match: 99.46% Indels: 0
 DB: 23 Gaps: 0

US-09-856-061-3 (1-1129) x AAE22609 (1-428)

OY 1 TTCGAGAACTTCAGTGTGCCAAAAAAGAGTCATGCGCTCGCATATAGTCCACAGGC 60
 DB 19 PheGlnAsnPheserLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGly 38
 OY 61 CAGTACCGAGAGATGAACAGCCCTCTCTAGACTGGGAAAGAACTTCTGCACTCTG 120
 DB 39 GlnTYrGlnArgMetAsnLysProLeuLeuAspTrpGlnArgAsnPhelAlaValLeu 58
 OY 121 GATGAGCAAAAGGCGCACAGTATGATGATGATGATGATGATGATGATGATGATG 180
 DB 59 AspGlyAlaLysGlyHisSerAspAspAspTrpAspAspProGlnLeuArgMetGlnGlu 78
 OY 161 ACATGGCAGTGCATTAATTTTACCAGCCGCGCTATTAAGAAATCTGAATATGACAT 240
 DB 79 ThrTrpGlnSerIleLysIleLeuProAlaArgProIleLysGlnSerGluTrpAlaAsp 98
 OY 241 ACACACTATTTCAAGCTTGCAATGACACTCCCTTCGCTTACAGACACGAGACTCTATC 300
 DB 99 ThrHisTYrPheLysValAlaMetAspThrProLeuProLeuAspThrArgThrSerIle 118
 OY 301 TCCATTGGACAGCCGACCTGGAACACAGAGAGAGTTGGAAGAGTGGCAACACCCATT 360
 DB 119 SerIleGlyGlnProThrTrpAsnThrGlnThrArgLeuGlnArgValAlaAspLysProIle 138
 OY 361 TCCAGGAGCGTCAGAGCCAAAAAATTAAGAGATGCAATCCGTAAAGAAAGCAAGATT 420
 DB 139 SerLysAspValArgSerGlnAsnIleLysGlyAspAlaSerValArgLysAsnLysIle 158
 OY 421 CCTTACCACTCTCTCGGCTCTCATACACTTCGAGAGAAATGACACCTTGGCCCT 480
 DB 159 ProLeuProProProArgProLeuIleThrLeuProLysLysTyrGlnProLeuProPro 178
 OY 481 GAGCGGAGAGAGAGAGCGACCTTATCTAGAGACACACTTCCCAAGTCCAGAGGA 540
 DB 179 GluProGlnSerSerArgProProLeuSerGlnArgHisThrPheProGlnValGlnArg 198
 OY 541 ATGCCAGTCAGATTAAGGACTTAAGGACTTAAGGAGTCTGAGAGCAGAAAAAGTTCT 600
 DB 199 MetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnAlaIleLysValPro 218
 OY 601 CATACACAGAGAGAGCGTAATCAACTCTCTGTTAGAAAACCAAAATATCTCAAGAGATT 660

DB 219 HisAsnGlnArgLysProGlnSerThrHisLeuGlnAsnGlnAsnThrGlnGluIle 238
 OY 661 CCACCTTGCCATTAGACAGTCTTCATTACAGACAGACAGACAGTGTGCAAAACAGAGAT 720
 DB 239 ProLeuAlaIleLysSerSerSerSerPheThrThrSerAsnHisSerValGlnAsnArgAsp 258
 OY 721 CATAGAGAGGCGATCAGCCCTGTTCTCTCAGAGATGCGACCTCCACAGCTGCAGC 780
 DB 259 HisArgGlyGlyMetGlnProCysSerProGlnArgCysGlnProProAlaSerCysSer 278
 OY 781 CCTCAGAAAATATATCTGCGCTATTAATACACAGCTGAGACACACTTCCCAAAAG 840
 DB 279 ProHisGlnAsnIleLeuProTyrLysTyrThrSerTrpArgProPheProLysArg 298
 OY 841 TCTGATAGAAGAGTGCAGACACATGAATGATGATGAGAAATACAGCCGAGGCA 900
 DB 299 SerAspArgLysAspValGlnHisAsnGlnTyrTrpIleGlyGlnTyrSerArgGlnAla 318
 OY 901 GTGGAAGAGCGCATTCATGAGAGAGAAACAAGATGCTAGTTCTTGCGAGATTGTTCC 960
 DB 319 ValGlnGlnAlaPheMetLysGlnLysAspGlySerPheLeuValArgAspCysSer 338
 OY 961 ACAAAATCCAGAGAGAGCCCTATGTTGGCTGTGTTTATGAGACAAAGCTTACAT 1020
 DB 339 ThrLysSerLysGlnGlnProTyrValLeuAlaValPheTyrGlnAsnLysValTyrAsn 358
 OY 1021 GTAAAAATCCGCTTCCCTGAGAGAGAAATCAGAGTTGCTCGGAGCAGACTCAGAGA 1080
 DB 359 ValLysIleArgPheLeuGlnArgAsnGlnGlnPheAlaLeuGlyThrGlyLeuArgGly 378
 OY 1081 GATGAGAACTTTCATTCATGAGAGACATCAGACATCAGACATCAGACAT 1128
 DB 379 AspGlnLysPheAspSerValGlnLysPheIleGlnHisTyrLysAsn 394

RESULT 3

AAE22608 standard; Protein; 443 AA.

AAE22608;

26-JUL-2002 (first entry)

Human MIST protein #1.

Human: mast cell immunoreceptor signal transducer; MIST; immune disorder;

lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;

inflammation; cancer; gene therapy; cytosolic; immunomodulatory.

Homo sapiens.

Key Location/Qualifiers

Modified-site 84 /label= Tyrosine-phosphorylation_binding_site

Modified-site 111 /label= Tyrosine-phosphorylation_binding_site

Domain 306..311 /label= SH3_binding_proline-rich_motif

Domain 324..407 /label= SH2_domain

WO200226986-A2.

04-APR-2002.

28-SEP-2001; 2001WO-US30593.

29-SEP-2000; 2000US-237030P.

(BRIM) BRISTOL-MYERS SQUIBB CO.

Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;

WPI, 2002-372126/40.

DR N-PSDB: AAD35800.
 XX
 PT New isolated mast cell immunoreceptor signal transducer polypeptide,
 PT useful for treating immune disorder involving hyperactivity of B- or
 PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
 PS
 PS Claim 11: Fig 3: 171pp: English.
 XX
 CC The present invention relates to novel mast cell immunoreceptor signal
 CC transducer (MIST) proteins and polynucleotides encoding such proteins.
 CC MIST sequences of the invention are useful for preventing, treating or
 CC ameliorating a medical condition in mammalian subject. They are useful
 CC for treating an immune disorder involving hyperactivity of B- or T-
 CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
 CC lymphoma, tumor or thymoma in a mammal. MIST sequences are useful
 CC as targets for therapeutic intervention in immune cell disorders and
 CC inflammatory indications, for diagnosis and/or screening of disorders
 CC or diseases associated with expression of MIST, for screening for
 CC antagonists or inhibitors of the interaction of MIST with cellular
 CC signalling components. They are used in assays that detect activation
 CC or induction of various B and T-cell-related neoplasms or cancers.
 CC Sequences of the invention are also used in gene therapy. The present
 CC sequence is human MIST protein.
 CC
 XX Sequence 443 AA:
 XX
 Alignment Scores:
 Pred. No.: 5 5e-195 Length: 443
 Score: 2009.00 Matches: 374
 Percent Similarity: 99.73% Conservative: 1
 Best Local Similarity: 99.47% Mismatches: 1
 Query Match: 99.46% Indels: 0
 DB: 23 Gaps: 0
 US-09-856-061-3 (1-1129) x AAE22608 (1-443)
 QY 1 TTCGAGAGCTTCAGTCCGCAAAAACAGAGTCATGCGCTCGCATCAATAGTCCACAGGC 60
 DB 34 PheGlnAsnPheserLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGly 53
 QY 61 CAGTACGAGAGGATGAAAGCCCTCTCTAGACTGGGAAGAACTTGGTCGACGTCTG 120
 DB 54 GlnTyrGlnArgMetAsnLysProLeuLeuAspTrpGlnArgAsnPheAlaAlaValLeu 73
 QY 121 GATGAGCAAAAGCCACAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 180
 DB 74 AspGlyAlaLysGlyLysSerAspAspAspAspAspAspAspAspAspAspAspAspAsp 93
 QY 181 ACATGAGCTGATTAATTTTACAGCCCGGCTTAAGAGATCTGAATATGACAGAT 240
 DB 94 ThrTrpGlnSerIleLysIleLeuProAlaArgProIleLysGlnSerGlnTyrAlaAsp 113
 QY 241 ACACACATTTTCAAGTTGCAATGACATCCCTTCCGTTAGACACAGAGCTCTATC 300
 DB 114 ThrHisTyrPheLysValAlaMetAspTrpProLeuProLeuAspTrpArgThrSerIle 133
 QY 301 TCCATTGAGAGCCGAGCTGGAACACACAGAGTTGGAAGAGTGGACAAACCAT 360
 DB 134 SerIleGlyGlnProTrpAsnThrGlnThrArgLeuGlnValAspLysProIle 153
 QY 361 TTCAGGAGCTCAGAACCCAAACATTAAGAGAGATCCCTTAAGAAAGACAGATT 420
 DB 154 SerLysAspValArgSerGlnAsnIleLysGlyAspAlaSerValArgLysAsnLysIle 173
 QY 421 CCTTACACCTCTCGGCTCTATACACTTCCGAAGAGTACCAACCTTCCCTCCCT 480
 DB 174 ProLeuProProProArgProLeuIleThrLeuProLysTyrGlnProLeuProPro 193
 QY 481 GAGCGGAGAGACAGCCACCTTATCTCAGAGACACACCTTCCAGAGAGTCCAGAGA 540
 DB 194 GlnProLysSerSerArgProProLeuSerGlnArgHisThrPheProGlnValGlnArg 213
 QY 541 ATGCCCAGTCAAGATTAAGGACTTAAGTGGAGTCTTGAAGCAGAAAAGTTCT 600

DB 214 MetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnAlaGlnLysValPro 233
 QY 601 CATACCAAGAGAGAGCTGAATCACTATCTGTAGAAAACCAAAATTAACAAGATT 660
 DB 234 HisAsnGlnArgLysProGlnSerThrHisLeuLeuGlnAsnGlnAsnThrGlnGlnIle 253
 QY 661 CCAGTTCGCAATTCAGCTTCTCATTCAGACAAAGCAACAGAGTGGCAACAGAT 720
 DB 254 ProLeuAlaIleSerSerSerSerPheThrThrSerHisSerValGlnAsnArgAsp 273
 QY 721 CATGAGAGAGCATGACAGCCCTGTTCTCTCAGAGATCCAGCCCTCCAGCCAGCTGCAGC 780
 DB 274 HisArgGlyGlyMetGlnProCysSerProGlnArgGlyGlnProAlaSerCysSer 293
 QY 781 CTCACGCAAAATATACCTCCCTATTAATTAACAAGCTGGAGACCACTTCCCAAAAG 840
 DB 294 ProHisGlnAsnIleLeuProTyrLysTyrThrSerTrpArgProProPheProLysArg 313
 QY 841 TCTGATGAAGAGATGTCACGACACAAATGATGATGAGAAATACAGCCGACAGCA 900
 DB 314 SerAspArgLysAspValGlnHisAsnGlnTrpTyrIleGlyGlnTyrSerArgGlnAla 333
 QY 901 GTGGAGAGGCATTCATGAAGAGAGACAGAGATGATGATGATGATGATGATGATGATGAT 960
 DB 334 ValGlnGlnAlaPheMetLysGlnAsnLysAspLysSerPheLeuValArgAspCysSer 353
 QY 961 ACAAAATCCAAAGAGAGCCCTATGTTTGGCTGTGTTTATGAGAACAAAGTCTACAAT 1020
 DB 354 ThrLysSerLysGlnGlnProTyrValLeuAlaValPheTyrGlnAsnLysValTyrAsn 373
 QY 1021 GTAAAAATCCGCTCTCTGAGAGAGAAATCAGACAGTTGCTCCGAGACAGACTCAGAGA 1080
 DB 374 ValLysIleArgPheLeuGlnArgAsnGlnGlnPheAlaLeuGlnTyrGlnLysLeuArgGly 393
 QY 1081 GATGAGAGTTTGAATTCATGAGACATCATGACATCATGACATCATGACATCATGACAT 1128
 DB 394 AspLysLysPheAspSerValGlnAspIleIleGlnHisTyrLysAsn 409
 RESULT 4
 ID AAE22610 standard; Protein: 353 AA.
 XX
 AC AAE22610:
 XX
 DT 26-JUL-2002 (first entry)
 XX
 DE Human MIST splice variant protein from clone #12.
 XX
 KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
 KW lymphocyte; lymphoma; tumor; thymoma; immune cell disorder; neoplasm;
 KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
 OS Homo sapiens.
 PN
 PN WO200226986-A2.
 PD 04-APR-2002.
 XX
 PF 28-SEP-2001; 2001WO-US30593.
 PR 29-SEP-2000; 2000US-237030P.
 XX
 PA (BRIM) BRISTOL-MYERS SQUIBB CO.
 PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
 DR WPI; 2002-372126/40.
 DR N-PSDB: AAD35802.
 PT New isolated mast cell immunoreceptor signal transducer polypeptide,
 PT useful for treating immune disorder involving hyperactivity of B- or
 PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma


```

QY 1 TTCAGAACTTCACTGCGCAAAAAACAGCTCATGCGCTCGCATCAATAGTCCACAGC 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19 PheGlnAsnValSerLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 38
QY 61 CAGTACCAAGAGATGAAACAGCTCTTCTAGATGCGAAAGAAACTTGTCTGACGCTCTG 120
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 39 ArgCysArgValValLeuGlnProLeuProAspHisArgAsnLeuValGlnValPro 58
QY 121 GATGAGCAAAAGCGCACATGATGATGATGATGATGATGATGATGATGATGATGATG 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59 G1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1y 78
QY 181 ACATGCGATCGATTAAATTTTACAGCCGCGCTATTAAGAAATCTGAATATGACAT 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 AlaTrpProSerMetLeuLeuLeuProAlaArgProLeuGlnGlnSerGlnValAla 98
QY 241 ACACACTATTTCAGATGCAATGACACATCCCTCCGTTAGACACAGACCTCTATC 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 ThrArgTrpPheGlnAspMetMetGlnAlaProLeuLeuLeuProProLysAlaSerVal 118
QY 301 TCCATTGGACAGCCGCTGGAAACACAG--ACGAGTTGGAAAGATGACAAACCC 357
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 119 SerThrGlnArgGlnThrArgAspValAlaArgMetThrGlnLeuGlnGlnValAspLysPro 138
QY 358 ATTTCAGGAGCAGTCAGAAAGCCAAACATTAAGAGATGATCCGTAGAAAGAAACAAG 417
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 139 ThrPheLysAspValArgSerGlnArgPheLysGlnPheLysTrpLysLysLysLys 158
QY 418 ATTCCTTACACACCTCCCTCCGCTCTCATACATTCGGAAGAGATCCAAACCTTGGCC 477
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 ThrProLeuProProProAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 178
QY 478 CCT--GAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 ProAlaProProGlnGlnSerSerLeuArgPheAlaProLysProThrPheProGlnVal 198
QY 535 CAGGAAATGCCAGTCAGATTAAGCTTAAAGGACTTAAGTGAAGTCTTGAAGCAGAAAA 594
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 GlnArgGlnProArgGlnArgSerLeuAlaLysAspPheSerArgValLeuGlnGln 218
QY 595 GTTCCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GluSerHisHisGlnThrLysProGlnSerSerLysProSerSerArgGlnAsnThrGln 238
QY 655 GAGATTCACCTTGCATTAAGCAGTCTTCTCATTCAGACAGAACACACAGTGTGCAAAAC 714
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 239 LysSerProProAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 258
QY 715 AGAGATCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 ArgAspHisThrGlnSerMetGlnHisCysProAlaGlnArgCysGlnAlaAlaLaser 278
QY 775 TGCAGCCCTCAGAAATATATACGCCCTATTAATACAAACACTGAGACCACTTTCCC 834
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 HisSerPro-----ArgMetLeuProTrpGlnAsnThrAsnSerGlnLysProAspPro 296
QY 835 AAAAGTCTGATAGAAAGATGTCACAGCACAATGAATGATGATGAGATGATGAGAGCCG 894
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 297 ThrLysProAspGlnLysAspValTrpGlnAsnGlnTrpLysLysLysLysLys 316
QY 895 CAGCAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 GlnAlaValGlnAspValLeuMetLysGlnLysAspGlnTrpPheLeuValArgAsp 336
QY 955 TGTTCACAAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 337 CysSerThrLysSerLysValGlnProTrpValLeuValAlaPheTrpGlnLysVal 356
QY 1015 TACAATGTAAATAATCCCTTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 357 TyrAsnValLysIleArgPheLeuGlnSerAsnGlnGlnPheAlaLeuGlnGlnThrGln 376

```

```

QY 1075 AGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1122
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 377 ArgGlnAsnGlnMetPheAspSerValGlnAspIleIleGlnHisLysTrp 392
RESULT 6
AACG03152
AACG03152 standard; Protein; 84 AA.
AACG03152:
06-OCT-2000 (first entry)
DE Human secreted protein, SEQ ID NO: 7233.
XX Human, 5' EST, expressed sequence tag; secreted protein; cDNA isolation;
XX gene therapy; chromosome mapping.
OS Homo sapiens.
PN EP1033401-A2.
PD 06-SEP-2000.
PF 21-FEB-2000; 2000EP-0200610.
PR 26-FEB-1999; 99US-0122487.
XX (GEST ) GENSET.
PA Dunas Milne Edwards J, Duclert A, Giordano J;
XX WPI: 2000-500381/45.
XX N-PSDB: AAC03158.
XX
XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for
XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
XX diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX Claim 13; SEQ ID 7233; 71pp + CD-ROM; English.
XX
XX The present sequence is a polypeptide encoded by one of a large number
XX of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
XX were prepared from total human RNAs or polyA+ RNAs derived from 30
XX different tissues. EST sequences usually correspond mainly to the 3'
XX untranslated region (UTR) of the mRNA because they are often obtained
XX from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
XX isolating cDNA sequences derived from the 5' ends of mRNAs and even in
XX those cases where longer cDNA sequences have been obtained, the full 5'
XX UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
XX ends and can therefore be used to obtain full length cDNAs and genomic
XX DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
XX chromosome mapping procedures. They are used to obtain upstream
XX regulatory sequences and to design expression and secretion vectors.
XX
XX Sequence 84 AA:
SQ
Alignment Scores:
Pred. No.: 1 02e-26 Length: 84
Score: 349.00 Matches: 64
Percent Similarity: 96.97% Conservative: 0
Best Local Similarity: 96.97% Mismatches: 2
Query Match: 17.28% Indels: 0
DB: 21 Gaps: 0
US-09-856-061-3 (1-1129) x AACG03152 (1-84)
QY 1 TTCAGAACTTCACTGCGCAAAAAACAGGTCATGCGCTCGCATCAATAGTCCACAGC 60
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 19 PheGlnAsnPheSerLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGly 38
QY 61 CAGTACCAAGAGATGAAACAGCTCTTCTAGATGCGAAAGAAACTTGTCTGACGCTCTG 120
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 39 ArgCysArgValValLeuGlnProLeuProAspHisArgAsnLeuValGlnValPro 58
QY 121 GATGAGCAAAAGCGCACATGATGATGATGATGATGATGATGATGATGATGATGATG 180
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 59 G1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1yG1y 78
QY 181 ACATGCGATCGATTAAATTTTACAGCCGCGCTATTAAGAAATCTGAATATGACAT 240
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 79 AlaTrpProSerMetLeuLeuLeuProAlaArgProLeuGlnGlnSerGlnValAla 98
QY 241 ACACACTATTTCAGATGCAATGACACATCCCTCCGTTAGACACAGACCTCTATC 300
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 99 ThrArgTrpPheGlnAspMetMetGlnAlaProLeuLeuLeuProProLysAlaSerVal 118
QY 301 TCCATTGGACAGCCGCTGGAAACACAG--ACGAGTTGGAAAGATGACAAACCC 357
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 119 SerThrGlnArgGlnThrArgAspValAlaArgMetThrGlnLeuGlnGlnValAspLysPro 138
QY 358 ATTTCAGGAGCAGTCAGAAAGCCAAACATTAAGAGATGATCCGTAGAAAGAAACAAG 417
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 139 ThrPheLysAspValArgSerGlnArgPheLysGlnPheLysTrpLysLysLysLys 158
QY 418 ATTCCTTACACACCTCCCTCCGCTCTCATACATTCGGAAGAGATCCAAACCTTGGCC 477
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 159 ThrProLeuProProProAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 178
QY 478 CCT--GAGCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 534
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 179 ProAlaProProGlnGlnSerSerLeuArgPheAlaProLysProThrPheProGlnVal 198
QY 535 CAGGAAATGCCAGTCAGATTAAGCTTAAAGGACTTAAGTGAAGTCTTGAAGCAGAAAA 594
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 199 GlnArgGlnProArgGlnArgSerLeuAlaLysAspPheSerArgValLeuGlnGln 218
QY 595 GTTCCTCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 654
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 219 GluSerHisHisGlnThrLysProGlnSerSerLysProSerSerArgGlnAsnThrGln 238
QY 655 GAGATTCACCTTGCATTAAGCAGTCTTCTCATTCAGACAGAACACACAGTGTGCAAAAC 714
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 239 LysSerProProAlaLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 258
QY 715 AGAGATCTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 259 ArgAspHisThrGlnSerMetGlnHisCysProAlaGlnArgCysGlnAlaAlaLaser 278
QY 775 TGCAGCCCTCAGAAATATATACGCCCTATTAATACAAACACTGAGACCACTTTCCC 834
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 279 HisSerPro-----ArgMetLeuProTrpGlnAsnThrAsnSerGlnLysProAspPro 296
QY 835 AAAAGTCTGATAGAAAGATGTCACAGCACAATGAATGATGATGAGATGATGAGAGCCG 894
   :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 297 ThrLysProAspGlnLysAspValTrpGlnAsnGlnTrpLysLysLysLysLys 316
QY 895 CAGCAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 954
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 GlnAlaValGlnAspValLeuMetLysGlnLysAspGlnTrpPheLeuValArgAsp 336
QY 955 TGTTCACAAATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1014
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 337 CysSerThrLysSerLysValGlnProTrpValLeuValAlaPheTrpGlnLysVal 356
QY 1015 TACAATGTAAATAATCCCTTCTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1074
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 357 TyrAsnValLysIleArgPheLeuGlnSerAsnGlnGlnPheAlaLeuGlnGlnThrGln 376

```



```

Db 286 sArgIngluaIaValaGlSerProValaPheProProAlaGlnIleHisGlnly 306
QY 792 TATACGCCCTAT---AAATACACA---AGCTGGAGACCA-----CCTTCCC 833
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 306 sProIleProLeuProAlaGpPheThgluGlnIleGlnProThValaSpIlyProLeuPr 326
QY 834 CAAAGGCTGTGATGAAGAAGATGTCACGACAAATGA-----TG 872
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 326 oSerPheSerAsnSerThrlleSerIugIngluaIleValleuCysIysProTr 346
QY 873 GTACATTTGAGATACAGCCGACAGTGGAGAGGATTCATTCATGAGAGACAAGA 932
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 346 pTyraIaGlnIaIaCysAspArgIySerIaGlnIaIaLeuHisArgSerIuIySas 366
QY 933 TGGTATTTCTTGGTCCGAGATTTCCACAAAATCCAGAGAACCCCTATTTTGGC 992
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 366 pGlySerPheLeuIleArgIySerSerGlnHisAspSerIyGlnProTyThrlleuVa 386
QY 993 TGTGTTTTTNGACAAAGCTACAAATGTAATAATCCGCTCTGGAGAGAAATGACGA 1052
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 386 ValPhePheAsnIySargIaIyTrasnIleProValaIArgPheIleGlnIaThrllySgl 406
QY 1053 GTTGGCTTGGGAGAGACTCAGAGAGATGAGAAGTTTGATTCAGTACAGACATCAT 1112
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 406 nTyraIaIeGlnIyArgIyIySasnIyGlnIuIyTrPheGlySerValaIaGlnIleI 426
QY 1113 CGAACACTACAGAAAT 1128
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 426 eArgAsnHisGlnHis 431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
AAW59865
ID AAW59865 standard; Protein; 456 AA.
AC AAW59865;
XX
DT 20-NOV-1998 (first entry)
XX
DE Amino acid sequence of the human BLNK-1 protein.
XX
KW Human; BLNK-1; B cells linker protein-1; apoptosis; Grb2; PLC-gamma;
KW SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; Sos;
KW ras pathway; GDP; GTP; calcium pathway; antigen.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc-difference 327 /note="encoded by AGC"
FT XX
PN WC9833852-A1.
XX
PD 30-JUL-1998.
XX
PF 23-JAN-1998; 98MO-US01394.
XX
PR 17-MAR-1997; 97US-0819013.
PR 24-JAN-1997; 97US-0788322.
XX
PA (UNIV ) UNIV WASHINGTON.
XX
PI Chan AC, Fu C;
XX
DR WPI; 1998-427948/36.
XX
DR N-PSDB; AAV41901.
XX
PT Human B cell linker proteins - useful in the treatment of diseases
XX
PS Involving increased or decreased apoptosis
XX
PS Claim 4; Fig 1; 56pp; English.
XX
CC This is the amino acid sequence of the human BLNK-1 (B cells linker
CC protein-1) protein, used the treatment of diseases involving the

```

```

CC increase and decrease of apoptosis. in the method of the invention.
CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
CC intracellular calcium levels and Grb2 is an adapter molecule
CC containing two SH3 domains that mediate its interaction with the
CC guanine nucleotide exchange factor, Son of Sevenless (Sos) which in
CC turn activates the ras pathway by facilitating the exchange of GDP for
CC GTP on the ras molecule. Activation of both ras and calcium pathways
CC are required for efficient B cell antigen receptor function. BLNK
CC binding proteins can be used to identify BLNK proteins in a target
CC sample.
XX
SQ sequence 456 AA:
XX
Alignment Scores:
Pred. No.: 3,61e-20 Length: 456
Score: 287.50 Matches: 107
Percent Similarity: 38.50% Conservative: 57
Best Local Similarity: 25.12% Mismatches: 142
Query Match: 14.23% Indels: 120
DB: 19 Gaps: 18

US-09-856-061-3 (1-1129) x AAW59865 (1-456)
QY 60 CCAGTACAGAGGATGACAAAGCCCTCTTACAGCTGGAAAGAACTT----- 108
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 57 ProIaAspIuGlnIuGlnIuGlnIuTrpSerAsp-AspPheAspSerAspTyGlnuSnpProAs 76
QY 109 -----GCTCAGCTCGTGGATGGAGCAAAAGCCACAGTATATGACTA 152
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 76 pGlnHisSerAspSerGlnuMetTyraIaIleProAlaGlnIuGlnuIaAspAspSerTy 96
QY 153 TGATGACCCCTGAGCTTGGATGGAGAGACATGCGATGATTAATAATTTCACAGCCG 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 96 rGlnProProProValaGlnIuGlnu-----ThnArgProValaHisProIaIe 112
QY 213 GCCTATAAGAAATCTGAATATGACATACACACTATTTCAGGTGCAATGACACTCC 272
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 112 uProPheAlaIargIyGlnuTy----- 119
QY 273 CCTCCCTTAGACACACAGACCCCTATCTCATTTGACACGCCGACCTGGAACACAGAC 332
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 120 -----IleAspSrnArgSer----- 125
QY 333 GAGGTTGAAAGAGTGGACAAACCCATTCCAGAGAGCTGACAGACCAAAATTAAGG 392
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 126 -----GlnArgHisSerProPheSerIyThrlleuProSerIyS----- 139
QY 393 AGATGATCCGTAAAGAAAGAAAGATTCCTTTACACACTCTCGCTCATACACT 452
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 140 -----ProSerTrpProSerGlnuIySalaIArgIeThrSerThrlleuProIaIeThrla 158
QY 453 TCCGAAAGATACCAACCCCTGCCCCGAGCCG----- 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 158 AleuGlnIySProGln---ValProProIySProIySgIyLeuLeuGlnuSgIuIaIaAs 177
QY 487 -----GAGAGCAGCAG 497
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 pTyraValaProValaGlnuSasPnaSgIuAsnTyIleHisProThrIuSerIe 197
QY 498 GCCA-----CCTTATCTCAGAGACACACCTTTCAGAGTCCAG----- 537
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 197 rProProProGlnuIySalaPrometValaIaArgSerThrllySProAsnSerIeThrPr 217
QY 538 -----GGAATGCCAGTGCAGATTAAGGACTTAAGGAGTCCCTTGAAGCAGAAA 593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 217 oAlaSerProProGlnThrlaIaSerGlnIyArgAsn---SerIyIaIaTrpGlnuTrIySe 236
QY 594 A-----GTTCTCATACACAGAGAGCCT----- 618
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 236 rProProProIaIaIaProSerProLeuProArgAlaGlnIySlySProThrlProIe 256
QY 619 -GAATCAACTCATCTGTAGAAAACAAATAACTCAAGAGATTCCACTTGCATTAGCAG 677
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 256 ulyssthrthProvalAlaserInglInsn----- 266
QY 678 TTCTCATTCAGACAGAACACAGTGGCAAAACAGATCATAGAGAGC----- 732
Db 267 -AlaserSerValCysGluCylulysProleProAlaGluAqHlsArgGlySerSerHl 286
QY 733 -ATGACGCCCTGTTCTCTCAGAGATGCCAGCCTCCAGCAGCTGCAGCCCTGCAAAA 791
Db 286 sargIngluAlaValGlnserProvalPheProProAlaGlnLysGlnLehlsInly 306
QY 792 TTNACGCCCTT---AAATACACA---AGCTGGACACA-----CCTTCCC 833
Db 306 sProleProleuProArPhePheThrCylulysAsnProthValAspGlyProleuPr 326
QY 834 CAAAGCTGTGATAGAAAGATGTCCAGCACATGAA-----TG 872
Db 326 oLlePheSerSerAsnSerThrIleSerGluIngluAlaGlyValLeuCysLysProTr 346
QY 873 GTACATTCGAGAAATACAGCCCGCAGCAGTGAAGAGCATTCTATGAGAGACAAGA 932
Db 346 pTyralaGlyAlaCysAspArgLysSerAlaGluAlaLeuHlsArgSerAsnLysAs 366
QY 933 TGTAGTCTTCTGTCGCCGATGTTGTCACAAATCCAGAGAAAGCCCTATGTTGGC 992
Db 366 pGlySerPheLeuIleArgLysSerSerLysHisAspSerLysGlnProTyThrLeuVa 386
QY 993 TGTGTTTATGAGAACAAAGTCTACATGTAAATCCGCTTCTGAGAGAGATCAGCA 1052
Db 386 lValPhePheAsnLysArgValAlaThrIleProvalArgPheIleGluAlaThrLysGl 406
QY 1053 GTTTCGCCCGGGGACAGAGCTCAGAGAGATGAGAACTTGATTGATGAGAACATCAT 1112
Db 406 nTyralaLeuGlyArgLysLysAsnGlyGlnLutyrPheGlySerValAlaGluIleI 426
QY 1113 CGAACCTACAGAAAT 1128
Db 426 eArgAsnHlsGlnHls 431
RESULT 10
AAM59867
ID AAM59867 standard; Protein: 457 AA.
AC AAM59867;
XX AAM59867;
XX 20-NOV-1998 (first entry)
DE Amino acid sequence of the mouse BLNK protein.
XX
XX Mouse; BLNK: B cells linker protein; apoptosis; Grb2; PLC-gamma;
XX SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; Sos;
XX ras pathway; GDP; GTP; calcium pathway; antigen.
XX
XX Mus sp.
XX
XX W09832852-A1.
XX
XX 30-JUL-1998.
XX
XX 23-JAN-1998; 98WO-US01394.
XX
XX 17-MAR-1997; 97US-0819013.
XX 24-JAN-1997; 97US-0788322.
XX
XX (UNIW ) UNIV WASHINGTON.
XX
XX Chan AC, Fu C;
XX
XX WPL; 1998-427948/36.
XX N-PSDB; AAV41903.
XX
XX Human B cell linker proteins - useful in the treatment of diseases
XX PT involving increased or decreased apoptosis

```

```

PS Disclosure: Fig 5; 56pp; English.
XX
CC This is the amino acid sequence of the murine BLNK (B cells linker
CC protein) protein, used the treatment of diseases involving the
CC increase and decrease of apoptosis, in the method of the invention.
CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
CC intracellular calcium levels and Grb2 is an adapter molecule
CC containing two SH3 domains that mediate its interaction with the
CC guanine nucleotide exchange factor, Son of Sevenless (Sos) which in
CC turn activates the ras pathway by facilitating the exchange of GDP for
CC GTP on the ras molecule. Activation of both ras and calcium pathways
CC are required for efficient B cell antigen receptor function. BLNK
CC binding proteins can be used to identify BLNK proteins in a target
CC sample.
XX
SQ Sequence 457 AA:
XX
Alignment Scores:
Pred. No.: 1.07e-18 Length: 457
Score: 273.00 Matches: 96
Percent Similarity: 38.71% Conserved: 60
Best Local Similarity: 23.82% Mismatches: 155
Query Match: 13.51% Indels: 92
DB: 19 Gaps: 16
XX
US-09-856-061-3 (1-1129) x AAM59867 (1-457)
QY 94 TGGGAAAGAACTTGTCTGCGACCTCTGATGAGCAAAAGCCACAGT----- 141
Db 64 TTPSerAspAspPheAspSerAspTyrGlnAsnProAspGlnHlsArgSerGlnMet 83
QY 142 -----GATGATGACTGATGATGATGATGATGATGATGATGATGATGATGATG 174
Db 84 TyralMetProAlaGluGluThrGlyAspAspSerTyrGluProProAlaGluGln 103
QY 175 GAAGAGACATGCGACCTGATTAATTTTACCAGCCCGCCTATTAAGAACTGATAT 234
Db 104 -----GlnThrArgValAlaHlsProAlaLeuProPheThrArgGlyGlyLyr 119
QY 235 GCAGATACACACTTTTCAAGTTGCAATGGACACTCCCTCCGTTAGACACACAGACC 294
Db 120 ValAspAsnArg-----SerSerGlnArgHlsSerProProPheSerLysThr 135
QY 295 TCTATCTCCATTGGACAGCCGACCTGGAACACAGACAGAGTTGGAAGAGTGACAAA 354
Db 136 LeuProSer-----LysProSerTyr----- 142
QY 355 CCCATTCCAGAGAGCGTCAGAACCCAAACATTAAAGAGATGCATCCGTAAGAAGAAC 414
Db 143 ProSerAlaLysAlaArgLeuAlaSerThrLeuProAlaProAsnSerLeuGlnLysPro 162
QY 415 AAGATTCTTACCACTCCCTCGCTCTCAT----- 447
Db 163 GlnValPro-----ProLysProLysAspLeuLeuGlnAspGluAlaAspTyrValValPro 181
QY 448 -----ACACTTCCGAGAAGATACCAACCCCTGCCCCCTGAG 483
Db 182 ValGluAspAsnAspGluAsnTyrIleHlsProArgLysSerSerProProAlaGlu 201
QY 484 CCG-----GAGAGCAGCAGCCACCTTATCTCGAAGACACACACTTTCCA 528
Db 202 LysAlaProMetValAsnArgSerThrLysProAsnSerSerLysHlsMetSerPro 221
QY 529 -----GAGTCCAGGAGATGCCAGTCATAGTAAAGGACTTAAGAGCTTACGCTCT 582
Db 222 ProGlyThrValAlaGlyArgAsnSerCylValTyrAspSerLysSerSerLeuProAla 241
QY 583 GAAGCAGAAAAAGTTCTCTATACAGAGAGAGCCGATCAATCACTCATCTGTAGAAAC 642
Db 242 AlaProSerProleuProAlaGlyLysLysPro---AlaThrProleu----- 257
QY 643 CAAATATCTCAAGAGATGTCCACTTTCATAGCAGTTCTTCTTACGACGACACACAC 702

```

```

Db 258 ---LysThrThrProValPro---ProLeuProAsnAlaSerAsnValCysGluGluLys 275
QY 703 AGTGTGCAAAACAGATCATAGAGAGCC-----ATGCCAGCCCTGTTCTCTCAGANA 756
Db 276 ProValProAlaGluArgHisArgLysSerHisArgGlnAspThrValAlaInsPro 295
QY 757 TGCCAGCCCTCCAGCCAGCCAGCCAGCCAGCAAAATATATGCTCCCTATGAATACACAGC 816
Db 296 ValPheProPheThrGlnLysProValHisGlnLysProValProLeu----- 311
QY 817 TGGAGACCACTTTCCTCCCAAGG----- 840
Db 312 -----ProArgPheProGlnAlaGlySerProAlaAlaAspGlyProPheHisSerPhe 329
QY 841 -----TCTGATAGGAAG---GATGTCCACACACAAATGATGTCAT 879
Db 330 ProPheAsnLeuThrPheAlaAspGlnGluGlyLeuLeuGlyLysProTrrPylAla 349
QY 880 GGAGATACAGCCGCCAGCCAGTGGAGAGGCAATTCATGAGAGCAAGATGCTAGT 939
Db 350 GAlaLacCysAspArgLysPheAlaGluGlnAlaLeuHisArgSerAsnLysAspGlySer 369
QY 940 TTCTTGTCGCGAGATGTTCCACAAATCCAAAGAGACCCCTATGCTTTGCTGCTTT 999
Db 370 PheLeuIleArgLysSerPheGlnHisAspSerLysGlnProTyrThrLeuValAlaPhe 389
QY 1000 TATGAGAACAAAGTCTACAAATGTAAATAATCCGTTCTGAGAGCAAGATGTCG 1059
Db 390 PheAsnLysArgValAlaGlnLysPheValArgPheIleGlnAlaThrLysGlnTyrAla 409
QY 1060 CTGGGACAGAGCACTCAGAGAGATGAGATTTGATTGATAGACATCATCATGAAAC 1119
Db 410 LeuGlyLysLysLysAsnGlyGluGlyTyrPheGlySerValValGluIleValAsnSer 429
QY 1120 TACAGCAAT 1128
Db 430 HisGlnHis 432

RESULT 11
ABB38147
ID ABB38147 standard; Peptide; 46 AA.
XX
XX ABB38147;
XX
XX AC
XX
XX DT 04-FEB-2002 (first entry)
XX
XX DE Peptide #5653 encoded by human foetal liver single exon probe.
XX
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX OS Homo sapiens.
XX
XX PN WO200157277-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001MO-US00669.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX
XX PR 26-MAY-2000; 2000US-0207456.
XX
XX PR 30-JUN-2000; 2000US-0608408.
XX
XX PR 03-AUG-2000; 2000US-0632366.
XX
XX PR 21-SEP-2000; 2000US-0234687.
XX
XX PR 27-SEP-2000; 2000US-0236359.
XX
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR.
XX
XX DR WPI; 2001-483447/52.
XX
XX PT Human genome-derived single exon nucleic acid probes useful for

```

```

PT analyzing gene expression in human fetal liver -
XX
XX PS Claim 27; SEQ ID NO 30782; 633pp + sequence listing; English.
XX
XX CC The invention relates to a single exon nucleic acid probe for
XX CC measuring human gene expression in a sample derived from human foetal
XX CC liver. The single exon nucleic acid probes may be used for predicting,
XX CC measuring and displaying gene expression in samples derived from human
XX CC fetal liver. The present sequence is a peptide encoded by a single exon
XX CC nucleic acid probe of the invention.
XX CC Note: The sequence data for this patent did not form part of the
XX CC printed specification, but was obtained in electronic format directly
XX CC from WIPO at ftp.wipo.int/pub/published_pcl_sequences.
XX
XX SQ Sequence 46 AA:
XX
XX Alignment Scores:
XX Pred. No.: 1.49e-15 Length: 46
XX Score: 238.00 Matches: 46
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 11.78% Indels: 0
XX DB: Gaps: 0
XX
XX US-09-856-061-3 (1-1129) x ABB38147 (1-46)
XX
XX QY 931 GATGTAGATTTCCTGTCGAGATGTTCCACAAATCCAAAGAGAGCCCTATGTTTG 990
XX 1 AspGlySerPheLeuValArgAspCysSerThrLysSerLysGluLysProTyrValLeu 20
XX
XX QY 991 GCTGTCTTTTATGAGAACAAAGTCTACAAATGTAAATAATCCGTTCTCGAGAGCAATCAG 1050
XX 21 AlaValPheTyrClnAsnLysValTyrAsnValLysIleArgPheLeuGlnArgAsnGln 40
XX
XX QY 1051 CAGTTGCCCTGGGAGCA 1068
XX
XX DB 41 GlnPheAlaLeuGlyThr 46

RESULT 12
ABB23354
ID ABB23354 standard; Protein; 46 AA.
XX
XX AC ABB23354;
XX
XX DT 23-JAN-2002 (first entry)
XX
XX DE Protein #5353 encoded by probe for measuring heart cell gene expression.
XX
XX KW Human; gene expression; heart; microarray; vascular system;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease.
XX
XX OS Homo sapiens.
XX
XX PN WO200157274-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001MO-US00666.
XX
XX PR 04-FEB-2000; 2000US-0180312.
XX
XX PR 26-MAY-2000; 2000US-0207456.
XX
XX PR 30-JUN-2000; 2000US-0608408.
XX
XX PR 03-AUG-2000; 2000US-0632366.
XX
XX PR 21-SEP-2000; 2000US-0234687.
XX
XX PR 27-SEP-2000; 2000US-0236359.
XX
XX PR 04-OCT-2000; 2000GB-0024263.
XX
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR.
XX
XX DR WPI; 2001-488899/53.
XX
XX PT

```

XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15: SEQ ID No 25124; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA2135-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 46 AA:
XX
Alignment Scores:
Pred. No.: 1,49e-15 Length: 46
Score: 238.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.78% Indels: 0
DB: 22 Gaps: 0
XX
US-09-856-061-3 (1-1129) x ABB23354 (1-46)
XX
QY 931 GATGCTAGTTCTTGTCGAGATTGTCACAAATCCAGAAGAGCCCTATGTTTG 990
DB 1 AspGlySerPheLeuValArgAspCysSerThrIlySerIySgluInuProTyValIeu 20
XX
QY 991 GCTGTGTTTATGAGCAAAAGTCTACATGTAAATCCGCTTCTCGAGAGAAATCAG 1050
DB 21 AlAlValPheTyrgIuAsnIySValTyfAsnValIySleArPheLeuGluArgAsnGln 40
XX
QY 1051 CAGTTTGCCTGGGAGACA 1068
DB 41 GlnPheAlaLeuGlyThr 46
XX
RESULT 13
AAM58783
ID AAM58783 standard; Protein; 46 AA.
XX
AC AAM58783;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 30888.
XX
KW Human: brain expressed exon; gene expression analysis; probe;
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN MO200157275-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX

PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4: SEQ ID NO: 30888; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 46 AA:
XX
Alignment Scores:
Pred. No.: 1,49e-15 Length: 46
Score: 238.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.78% Indels: 0
DB: 22 Gaps: 0
XX
US-09-856-061-3 (1-1129) x AAM58783 (1-46)
XX
QY 931 GATGCTAGTTCTTGTCGAGATTGTCACAAATCCAGAAGAGCCCTATGTTTG 990
DB 1 AspGlySerPheLeuValArgAspCysSerThrIlySerIySgluInuProTyValIeu 20
XX
QY 991 GCTGTGTTTATGAGCAAAAGTCTACATGTAAATCCGCTTCTCGAGAGAAATCAG 1050
DB 21 AlAlValPheTyrgIuAsnIySValTyfAsnValIySleArPheLeuGluArgAsnGln 40
XX
QY 1051 CAGTTTGCCTGGGAGACA 1068
DB 41 GlnPheAlaLeuGlyThr 46
XX
RESULT 14
AAM71296
ID AAM71296 standard; Protein; 46 AA.
XX
AC AAM71296;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31602.
XX
KW Human: bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; Leukemia; Lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN MO200157276-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001MO-US00668.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI: 2001-488900/53.
DR Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
PT
XX
PS Example 4: SEQ ID NO: 31602; 658pp + Sequence Listing: English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 46 AA;

Alignment Scores:
Pred. No.: 1.49e-15 Length: 46
Score: 238.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.78% Indels: 0
DB: 22 Gaps: 0

US-09-856-061-3 (1-1129) x AAM71296 (1-46)

QY 931 GATGGTAGTTCTGTGTCGAGATTGTTCCACAAATCCAGAGAGCCCTATGTTTG 990
DB 1 AspGlySerPheLeuValArgAspCysSerThrLysSerLysGluIuProTyValLeu 20
QY 991 GCTGCTTTATGAGAACAAAGCTACATGTAAATCCGCTTCTGGAGAGAAATCAG 1050
DB 21 AlaValPheTyGluAsnLysValTyArgAsnValLysIleArgPheLeuGluArgAsnGln 40
QY 1051 CAGTTGCCCTGGGACA 1068
DB 41 GlnPheAlaLeuGlyThr 46

RESULT 15
AAM18973
ID AAM18973 standard; Protein: 46 AA.
XX
AC AAM18973;
XX
DT 12-OCT-2001 (first entry)
XX
DE Peptide #5407 encoded by probe for measuring cervical gene expression.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
XX cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
PD
XX 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX

XX WPI: 2001-488901/53.
DR Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
PT
XX
PS Claim 27: SEQ ID NO 23799; 487pp; English.
XX
CC The present invention relates to human single exon nucleic acid probes
CC (SENP: see A110068-A1128459). The present sequence is a peptide encoded
CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 46 AA;

Alignment Scores:
Pred. No.: 1.49e-15 Length: 46
Score: 238.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.78% Indels: 0
DB: 22 Gaps: 0

US-09-856-061-3 (1-1129) x AAM18973 (1-46)

QY 931 GATGGTAGTTCTGTGTCGAGATTGTTCCACAAATCCAGAGAGCCCTATGTTTG 990
DB 1 AspGlySerPheLeuValArgAspCysSerThrLysSerLysGluIuProTyValLeu 20
QY 991 GCTGCTTTATGAGAACAAAGCTACATGTAAATCCGCTTCTGGAGAGAAATCAG 1050
DB 21 AlaValPheTyGluAsnLysValTyArgAsnValLysIleArgPheLeuGluArgAsnGln 40
QY 1051 CAGTTGCCCTGGGACA 1068
DB 41 GlnPheAlaLeuGlyThr 46

Search completed: April 21, 2003, 12:20:53
Job time : 58.706 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus.n2p model

Run on: April 21, 2003, 12:16:17 ; Search time 13.2707 Seconds

(without alignments)
5006.288 Million cell updates/sec

Title: US-09-856-061-3

Perfect score: 2020

Sequence: 1 ttccagacttcagctgcgc.....catgcacactacaagaatt.1129

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 262574 segs, 29422922 residues

Total number of hits satisfying chosen parameters: 525148

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODEL=frame+ n2p model -DEV=xlp
-Q/cg2n2_1/USPTO_SPOOL/US09856061/runat.21042003.113020.3774/app_query.fasta.1.3150
-DB=Issued_Patents_AA -QFMT=fastan -SUFIX=n2p.ra1 -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061.ecgn1.1.46.0runat.21042003.113020.3774 -NCPU=6 -ICPU=3
-NO_XIPXY -NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMOUT=120
-WARN_TIMOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELDP=6 -DELEXT=7

Database :

1: Issued_Patents_AA:*
2: /cg2n2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cg2n2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cg2n2_6/ptodata/1/1aa/5A.COMB.pep:*
5: /cg2n2_6/ptodata/1/1aa/5B.COMB.pep:*
6: /cg2n2_6/ptodata/1/1aa/5A.COMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	287.5	14.2	456	2	US-08-819-013-1
2	155	7.7	724	1	US-07-906-349A-5
3	155	7.7	724	1	US-08-167-035-2
4	155	7.7	724	1	US-08-208-887A-2
5	155	7.7	724	2	US-08-539-005-2
6	150.5	7.5	724	4	US-09-280-598-5
7	144.5	7.2	722	4	US-08-390-874C-12
8	144.5	7.2	722	4	US-09-265-772-12
9	143	7.1	659	1	US-08-391-615-2
10	140	6.9	659	4	US-08-426-509A-8
11	140	6.9	659	4	US-08-391-615-6
12	137	6.8	442	1	US-08-391-615-6

13	130	6.4	92	1	US-08-202-389-28	Sequence 28, App1
14	129.5	6.4	474	3	US-08-729-416C-1	Sequence 1, App11
15	129.5	6.4	594	3	US-08-729-416C-7	Sequence 7, App11
16	125	6.2	593	1	US-08-202-389-12	Sequence 12, App1
17	125	6.2	593	1	US-08-018-129-5	Sequence 5, App11
18	125	6.2	593	2	US-08-448-250-5	Sequence 5, App11
19	124.5	6.2	160	1	US-08-167-035-40	Sequence 40, App1
20	124.5	6.2	160	1	US-08-208-887A-40	Sequence 40, App1
21	124.5	6.2	160	2	US-08-539-005-40	Sequence 40, App1
22	124.5	6.2	160	4	US-09-280-598-37	Sequence 37, App1
23	123	6.1	454	4	US-08-259-264-2	Sequence 2, App11
24	121.5	6.0	1893	1	US-08-188-582-11	Sequence 11, App1
25	121.5	6.0	1893	1	US-08-646-715-11	Sequence 11, App1
26	120	5.9	694	3	US-08-559-397A-31	Sequence 31, App1
27	119	5.9	99	2	US-08-820-754-25	Sequence 25, App1
28	119	5.9	99	3	US-08-956-652-25	Sequence 25, App1
29	119	5.9	99	3	US-08-856-869-25	Sequence 25, App1
30	119	5.9	99	3	US-08-948-547-25	Sequence 25, App1
31	119	5.9	620	4	US-08-426-509A-9	Sequence 9, App11
32	119	5.9	620	5	PCT-US95-05008-9	Sequence 9, App11
33	119	5.9	1872	1	US-08-188-582-14	Sequence 14, App1
34	119	5.9	1872	1	US-08-646-715-14	Sequence 14, App1
35	117.5	5.8	474	3	US-08-807-342B-2	Sequence 2, App11
36	117	5.8	95	1	US-08-308-086-8	Sequence 8, App11
37	117	5.8	96	2	US-08-479-078-20	Sequence 20, App1
38	117	5.8	102	1	US-08-202-389-27	Sequence 27, App1
39	116	5.7	130	4	US-09-442-100-4	Sequence 4, App11
40	115	5.7	675	5	US-08-426-509A-4	Sequence 4, App11
41	115	5.7	675	5	PCT-US95-05008-4	Sequence 4, App11
42	114.5	5.7	625	1	US-08-391-615-3	Sequence 3, App11
43	113	5.6	226	3	US-09-195-868-30	Sequence 30, App1
44	113	5.6	1229	3	US-09-195-868-28	Sequence 28, App1
45	112	5.5	99	3	US-09-195-868-23	Sequence 23, App1

ALIGNMENTS

RESULT 1
US-08-819-013-1
Sequence 1, Application US/08819013
Patent No. 5994522
Parent Information:
Applicant: Chan, Andrew C.
Title of Invention: BLNK PROTEINS
Number of Sequences: 13
Correspondence Address:
Addressee: Flehr, Hohbach, Test, Albritton & Herbert
Street: Four Embarcadero Center, Suite 3400
City: San Francisco
State: California
Country: United States
Zip: 94111-4187
Computer Readable Form:
Medium Type: Floppy disk
Computer: IBM PC compatible
Operating System: PC-DOS/MS-DOS
Software: Patentin Release #1.0, Version #1.30
Current Application Data:
Application Number: US/08/819,013
Filing Date: 17-MAR-1997
Classification: 435
Prior Application Data:
Application Number: US 08/788,322
Filing Date: 24-JAN-1997
Attorney/Agent Information:
Name: Silva, Robin M.
Registration Number: 38,304
Reference/Docket Number: A-64383-1/RFT/RMS
Telecommunication Information:
Telephone: (415) 781-1989
Telefax: (415) 398-3249
Telex: 910 277299
Information for SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:
 LENGTH: 456 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 US-08-819-013-1

Alignment Scores:

Pred. No.:	2,72e-21	Length:	456
Score:	287.50	Matches:	107
Percent Similarity:	38.50%	Conservative:	57
Best Local Similarity:	25.12%	Mismatches:	142
Query Match:	14.23%	Indels:	120
DB:	2	Gaps:	18

US-09-856-061-3 (1-1129) x US-08-819-013-1 (1-456)

```

QY 60 CCAGTACGAGAGATGAACACGCTCTTCTAGACTGGGAAAGAACTTT----- 108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 57 ProAlaAspGluGluGlnGlnIlePheSerAsp-AspHeAspSerAspIYrGluAsnProAs 76
QY 109 -----GCTGCAGTCCCTGGAGTGACCAAAAGCCACAGTATGATGACTA 152
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 76 pEluHisSerAspSerGluMetIYr-ValMetProAlaGluGlnGluAsnAlaAspSerTy 96
QY 153 TGATGACCCCTGACCTTGGATGGAGAGACATGCGATTAATAATTTTCCAGCCCG 212
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 96 rGluProProProValGluGlnGlu-----ThrArgProValHisProAlaLe 112
QY 213 GCCTATTAAGAAATCTGATATGATGACATACACTATTTCAGAGTTCGAATGACACTCC 272
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 112 uProPheAlaArgGlyGlyTy----- 119
QY 273 CCTTCCGTTAGACACGACGACTCTATCTCCATTGGACACCGGACCTGGAAACACAGAC 332
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 120 -----IleAspAsnArgSerSer----- 125
QY 333 GAGGTTGGAAGAGTGGCAAAACCATTTCCAGGAGCGTCAGAGCCAAACATTAAGG 392
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 126 -----GlnArgHisSerProPheSerIYrThrIeuProSerIYs----- 139
QY 333 AGATGATCCCTAGAAAGAACAGATTCCTTACCACCTCTCGCCCTCATTAACACT 452
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 140 ----ProSerTrpProSerGluYsAlaArgIeuThrSerThrIeuProAlaLeuThrAl 158
QY 453 TCCGAAGAGATGACCAACCTTGCCTGAGCCG----- 486
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 158 aLeuGluYsProGln--ValProProYsProYsGlyIeuLeuGluAspGluAlaAs 177
QY 487 -----GAGAGCAGCAG 497
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 177 pTyrValValProValGluAspAsnAspGluAsnIYrIleHisProThrIeuSerSerSe 197
QY 498 GCCA-----CCTTATCTCAGAGACACACCTTTCAGAGTCAG----- 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 197 rProProProGluYsAlaProMetValAsnArgSerThrIYsProAsnSerSerThrPr 217
QY 538 ----GGAATGCCAGCTGATAGCTTAAGGACTTAAGTGAAGCTCTTGAAGCAGAA 593
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 217 oAlaSerProGluYrAlaSerGlyArgAsn--SerIYalatrPgluThrIYsSe 236
QY 594 A-----GTTCCTCATTAACGAGAGAGCT----- 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 236 rProProProAlaAlaProSerProLeuProArgAlaGlyIYsIYsProThrIYrProLe 256
QY 619 -GAATCAACTATCTGTTAGAAACCAAAATACTCAAGAGATTCACCTTGCATTAGCAG 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 256 uIYsThrThrProValAlaSerGlnIYAsn----- 266
QY 678 TCTTCATTCAGCAACGACACACAGTGTGCCAAACAGACATCATAGAGAGCC----- 732
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 267 -AlaSerSerValYsGluGlnIYsProIleProAlaGluArgHisArgGlySerSerHI 286
  
```

```

QY 733 -ATGAGCCCTGTCTCTCCACAGATGCCAGCCTCCAGCCACTCCAGCCCTCAGGAAA 791
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 286 sArgGlnGluAlaValGlnSerProValPheProProAlaGlnGlnIleHisGlnIY 306
QY 792 TATACTGCCCTTAT-----AATACACA---AGCTGAGAGACA-----CCTTCCC 833
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 306 sProIleProLeuProAlaGlyPheThrGluGlyIYsIYsProThrValAspGlyProLeuPr 326
QY 834 CAAAAGTCTGTAGAAAGATGTCCACACATGAA-----TG 872
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 326 oIlePheSerSerAsnSerThrIleSerGluGlnGluAlaGlyValLeuCYsIYsProTr 346
QY 873 GTACATTTGAGATACAGCCGCGCAGCAGTGGAAAGAGCATTCATCAGACGACAAACAGA 932
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 346 pTyrAlaGlyAlaCYsAspArgIYsSerAlaGluIYsIleuHisArgSerAsIYsAs 366
QY 933 TGTGATTTCTGTGTCGAGATTTGCCAAMAAATCCAAAGAGAGCCCTATGTTTGAGC 992
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 366 pOlySerPheLeuIleArgIYsSerSerGlyHisAspSerIYsGlnProTYrThrLeuVa 386
QY 993 TGTGTTTATGAGCAAAAGTCTACATGTAAATAATCCGCTTCTGAGAGAGATCACA 1052
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 386 lValPhePheAsnIYsArgValTYrAsnIleProValArgPheIleGluAlaThrIYsGl 406
QY 1053 GTTTCCTCGGGACAGACACTCAGAGAGATGAGAAAGTTGTTGATGAGAAAGATCAT 1112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 406 nTyrAlaLeuGlyArgIYsIYsAsnGlyIYsIYsPheGlySerValAlaGluIleIl 426
QY 1113 CGAACACTACAAAGAT 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 426 eArgAsnHisGlnHis 431

RESULT 2
US-07-906-349A-5
: Sequence 5, Application US/07906349A
: Patent No. 5434064
: GENERAL INFORMATION:
: APPLICANT: Schlessinger, Joseph
: APPLICANT: Skolnik, Edward Y.
: APPLICANT: Margolis, Benjamin I.
: TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
: TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES
: NUMBER OF SEQUENCES: 16
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Browdy and Neimark
: STREET: 419 Seventh Street, N.W.
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20004
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/07/906,349A
: FILING DATE: 30-JUN-1992
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 07/643,237
: FILING DATE: 18-JAN-1991
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 202-628-5197
: TELEFAX: 202-737-3528
: INFORMATION FOR SEQ ID NO: 5:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 724 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
  
```

MOLECULE TYPE: protein
US-07-906-349A-5

Alignment Scores:

Pred. No.: 1.79e-07 Length: 724
Score: 155.00 Matches: 94
Percent Similarity: 37.24% Conservative: 52
Best Local Similarity: 23.98% Mismatches: 103
Query Match: 7.67% Indels: 143
DB: 1 Gaps: 20

US-09-856-061-3 (1-1129) x US-07-906-349A-5 (1-724)

```

OY 308 GACAGCCAGCCAGCAACACAGAGAGTGGAAAGTGACAAC----- 355
DB 47 GIUAIATGPR-----ArgHrAsnGlyTrPLeuAsnGlyTrAsnGluThrGly 64
OY 356 -----CCATTTCAGGAGCGTCAAGAACATTAAAGAGATGATCCGTAA 406
DB 65 GIUYSGLYAspPheProGlyThr-----TyrValGluTyrIleGly-Argly 80
OY 407 GAAAGAACAAAGTTCCTTACACCTCGGCTCTCTACACTTCGAGAGAGTACC 466
DB 80 slyIleSerPro-ProThrProlysProAlqPro-----ProArg----- 93
OY 467 AACCTTGCCCGCCGAGCGAGAGCAGCAGG----- 498
DB 94 --ProleuProValAlaProGlySerSerLysThGluAlaAspValGluGlnAla 113
OY 499 -----CCACCTTATCTCAGACACACCTTCCAGAGTC----- 534
DB 113 euThLeuProAspLeuAlaGluGlnPheAlaProProAspLeuAlaProProLeu 133
OY 535 -----CAGGAATG-----C 544
DB 133 IeLysLeuValGluAlaIleGluLysLysGlyLeuGluLysSerThrLeuTyrArgThG 153
OY 545 CCAGTCAGATTAAGCTTAAGGAGCTTAAGAGAGCTTCCTGAGCAAGAAAGTTCTCAT 604
DB 153 InSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 172
OY 605 ACCAGAGAGAGCTGAGTCACTCATCTGTAGAAACCA----- 645
DB 172 alAspLeuLueLleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuLeu 192
OY 646 -----AATCTCAAGAGATTCACCTGCAATGAGCTCT----- 681
DB 192 sPLeuProAsnProValIleProAlaAlaValTyrSerGluMetIleSerLeuAlaPro 212
OY 681 ----- 681
DB 212 IuValGInSerSerGluGluTyrIleGlnLeuLysLysLeuLysLeuArgSerProSer 232
OY 682 -----TCAT 685
DB 232 IeProHISGlnTyrTrPLeuThrLeuGlnTyrLeuLysHisPhePheLysLeuSer 252
OY 686 TCACGCAAGCAACACAGTGTGCAAAACAGATCATAGAGAGCAGCAGCCCTGT 745
DB 252 InThSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 271
OY 746 CTCCTCAGA-----GATGACAGCTCCAGCCAGCTGACAGCCCTCAGAAA---ATATAC 796
DB 272 LeuPheArgPheSerAlaAlaSerSerAspAsnThGluAsnLeuIleLysValIleGlu 291
OY 797 TGCCCTATAATACACAGCTGAGA-----CCACCTTTC 832
DB 292 IleLeuIleSerThrGlu--TrpAsnGluArgGlnProAlaProAlaLeuProProLys 311
OY 833 CCAAAAG-----TCTGATTAAGAAAGATGTCAGACACA 865
DB 311 roProLysProThrThrValAlaAsnAsnGlyMetAsnAsnMetSerLeuGlnAsnA 331

```

```

OY 866 ATGATGTACATTGAGATAATACCCCGCAGTGGAGAGGCAATTCATGAGAGAGA 925
DB 331 IagIuTrpTyrTrpGlyAspIleSerArgGluGlnValAsnGluLys---LeuArgAsp 350
OY 926 ACAAGATAGTATTCTTGGTCCGAGATTGTTCCACAAATCCAGAGAGACCCCTATG 985
DB 350 hrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp---TyrT 369
OY 986 TTTGGCTGTGTTTATGAGAACAAAGCTCAATGTAAA---ATCCGCTTCCTGAGAA 1042
DB 369 hrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIlePheHis 386
OY 1043 GGAATCAGACAGTTTCCTCGGAGCAGACTCAGAGAGATGAGAGATTGATCAGTAG 1102
DB 386 rGAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheSerSerVal 402
OY 1103 AAGACATCATCGAACACTRCAAGAAAT 1128
DB 402 alGluLeuIleAsnHisTyrArgAsn 410

```

RESULT 3

US-08-167-035-2

Sequence 2, Application US/08167035

Patent No. 5618691

GENERAL INFORMATION:

APPLICANT: Schlessinger, Joseph

APPLICANT: Skolnick, Edward Y.

APPLICANT: Margolis, Benjamin L.

TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR

TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE

TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS

NUMBER OF SEQUENCES: 50

CORRESPONDENCE ADDRESS:

ADDRESSEE: PENNIE & EDMONDS

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: 10036-2711

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/167,035

FILING DATE: 16-DEC-1993

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A.

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7683-062

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

INFORMATION FOR SEQ. ID NO. 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 724 amino acids

TYPE: amino acid

TOPOLOGY: unknown

MOLECULE TYPE: protein

US-08-167-035-2

Alignment Scores:

Pred. No.: 1.79e-07 Length: 724

Score: 155.00 Matches: 94

Percent Similarity: 37.24% Conservative: 52

Best Local Similarity: 23.98% Mismatches: 103

Query Match: 7.67% Indels: 143

DB: 1 Gaps: 20

US-09-856-061-3 (1-1129) x US-08-167-035-2 (1-724)

```

QY 308 GACAGCCGACCTGGAACACACAGAGGTTGGAAGAGTGCAAC----- 355
Db 47 GLUALAARGPro-----ArgArgasnnglytripleuasnglytyrasmngluthrthrgly 64
QY 356 -----CCATTTCAGGAGAGCTGAGAACCCAAACATTAAAGAGATCGCTGA 406
Db 65 GLUysglYAspPheProglythr-----TyrValGluTyrilegly-Argly 80
QY 407 GAAAGAACAGATTCCTTACCACTCGGCTCTCATACACTTCGGAAGAGTACC 466
Db 80 slyslIleSerPro-ProthrProlysProArgPro-----ProArg----- 93
QY 467 AACCCCTGCCCCGAGCGAGAGACAGAG----- 498
Db 94 --ProleuProValAlaProGlySerSerlysthrGluAlaAspValGluGlnAlaL 113
QY 499 -----CCACCTTTATCTCAGAGACACACCTTTCAGAAAGTC----- 534
Db 113 eutThrleuProAspleuAlaGluGlnPheAlaProProAspleuAlaProProleuLeuI 133
QY 535 -----CAGCAATG-----C 544
Db 133 IeLysleuValGluAlaIleGluLysLysGlyLeuGluCysSerThrLeuTyrArgThrg 153
QY 545 CCAGTCAGATAGCTTAAAGGAGCTTAAGTGCCTTGAAGCAGAAAGTTCCTCAT 604
Db 153 InSerSerAsnleuAlaGluLeuArgGlnleuLeuAspCysasp---ThrProSerV 172
QY 605 AACGAGAGACCTGATCACTATCTGTTAGAAACCA----- 645
Db 172 alaAspleuGluMetIleAspValIleValleuAlaAspAlaPheLysArgTyrleuLeuA 192
QY 646 -----AATACTCAAGAGATTCACCTTGCCTTAGACGTTCT----- 681
Db 192 spleuProAsnProValIleProAlaAlaValItyrSerGluMetIleSerleuAlaProG 212
QY 681 ----- 681
Db 212 IuValGlnSerSerGluGluTyrIleGlnleuLeuLysLysleuIleArgSerProserI 232
QY 682 -----TCAT 685
Db 232 IeProHISGlnTyrTrpleuThrleuGlnTyrleuLeuLysHisPhePheLysleuSerG 252
QY 686 TCAGAGACAGAACCCAGAGTGTGCAAAAACAGATCATAGAGAGCATCGACCTGT 745
Db 252 InThrSerSerLysAsnleuLeuAsnAlaArg-ValleuSerGluIlePheSerProMet 271
QY 746 CTCCTCAG-----GATGCCAGCCTCAGACAGCTGACGACCCCTCAGAAA---ATATAC 796
Db 272 leuPheArgPheSerAlaAlaIleSerSerAspAsnThrGlnleuIleLysValIleGlu 291
QY 797 TGCCCTAATAATACACAGCTGAGAG-----CCACCTTCC 832
Db 292 IleleuIleSerThrGlu--TyrAsnGlnArgGlnProAlaProAlaLeuProProLysP 311
QY 833 CCAAAAG-----TCTGATGAAAGAGATGCTCCAGACA 865
Db 311 ropProLysProThrThrValAlaAsnAsnGlyMetAsnAsnAsnMetSerleuInAsnA 331
QY 866 ATGAATGTCATGTGAGATATACGCCGCGAGAGTGAAGAGGCAATTCATGAGAGAGA 925
Db 331 IeGluTyrPyrTrpGlyAspIleSerArgGlnGluValaIngLys---LeuArgAspR 350
QY 926 ACAAGATGAGTATCTTGTGCGAGATGTTCCACAAATCCAAAGAGAGAGCTATG 985
Db 350 hrAlaAspGlyThrPheleuValaIArgAspAlaSerThrIlysmethISglYAsp---TyrT 369
QY 986 TTTTGGCTGCTGTTTATGAGAACAACTGACAAATGTAA---ATCCGCTCTCTGAGAGA 1042
Db 369 hrleuThrleu-----ArgLysglYglYAsnAsnLysleuIleLysIlePheHisA 386

```

```

QY 1043 GGAATCAGACGTTTCCCTGGGAGACAGACTCAGAGAGATGAGAGATTGATTCAGTAG 1102
Db 366 rGAspGlyLysTyrIlePheSerAspProleu-----ThrPheSerValV 402
QY 1103 AAGACATCATCGAACATCAACAAGAT 1128
Db 402 alGluLeuIleAsnHisTyrArgAsn 410

RESULT 4
US-08-208-887A-2
: Sequence 2, Application US/08208887A
: Patent No. 5677421
: GENERAL INFORMATION:
: APPLICANT: Schlessinger, Joseph
: APPLICANT: Skolnick, Edward Y.
: APPLICANT: Margolis, Benjamin L.
: TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
: TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
: TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS
: NUMBER OF SEQUENCES: 51
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: PENNIE & EDMONDS
: STREET: 1155 Avenue of the Americas
: CITY: New York
: STATE: New York
: COUNTRY: 10036-2711
: ZIP: 10036-2711
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/208,887A
: FILING DATE: 11-MAR-1994
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Cornuzi, Laura A.
: REGISTRATION NUMBER: 30,742
: REFERENCE/DOCKET NUMBER: 7683-063
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (212) 790-9090
: TELEFAX: (212) 869-9741/8864
: INFORMATION FOR SPO ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 724 amino acids
: TYPE: amino acid
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: US-08-208-887A-2

Alignment Scores:
Pred. No.: 1.79e-07 Length: 724
Score: 155.00 Matches: 94
Percent Similarity: 37.24% Conservative: 52
Best Local Similarity: 23.98% Mismatches: 103
Query Match: 7.67% Indels: 143
DB: 1 Gaps: 20

US-09-856-061-3 (1-1129) x US-08-208-887A-2 (1-724)
QY 308 GACAGCCGACCTGGAACACACAGAGGTTGGAAGAGTGCAAC----- 355
Db 47 GLUALAARGPro-----ArgArgasnnglytripleuasnglytyrasmngluthrthrgly 64
QY 356 -----CCATTTCAGGAGAGCTGAGAACCCAAACATTAAAGAGATCGCTGA 406
Db 65 GLUysglYAspPheProglythr-----TyrValGluTyrilegly-Argly 80
QY 407 GAAAGAACAGATTCCTTACCACTCGGCTCTCATACACTTCGGAAGAGTACC 466
Db 80 slyslIleSerPro-ProthrProlysProArgPro-----ProArg----- 93

```

```

OY 467 AACCTTGCCTTGAGCCGAGACAGCAGG----- 498
    |||||  |||  |||||
Db 94 --ProleuproValaIalaproGlySerSerlysthrGluAlaaspValglungInGlnAlaI 113
OY 499 -----CCACCTTTATCTCAGACACACACCTTTCCAGAAATC----- 534
    |||  |||  |||  |||  |||  |||
Db 113 euthrleuproaspIleuAlaglInGlnpheAlaproProaspIleAlaproProleuLeuI 133
OY 535 -----CAGGGAATG-----C 544
    |||  |||  |||  |||
Db 133 IeLysLeuValgluAlaIleglLysLysGlyLeuGluCysSerThrLeuTyraThrG 153
OY 545 CCAGTCACATAGACTTTAAGGAGCTTAAGTCAAGCTTGAGCAGAGAAAGTTCCTCAT 604
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 153 InsetSerSerAsnLeuAlaIleuAlaArgInLeuLeuAspCysasp---ThrProSerV 172
OY 605 ACCAGAGAGAGCCTGAATCACTCATCTGTGTAGAAACCA----- 645
    |||  |||  |||  |||  |||  |||
Db 172 AlaSpLeuGluMetIleAspValHisValLeuAlaAspAlaPheLysArgTyrlLeuAla 192
OY 646 -----AATACTCAGAGATTCACCTGCCATTAGCAGTCT----- 681
    |||  |||  |||  |||  |||  |||
Db 192 spleuProaspProValIleProAlaAlaValTySerGluMetIleSerLeuAlaProG 212
OY 681 ----- 681
Db 212 LuValGlnSerSerGluGluTyrlleGlnLeuLeuLysLysLeuIleArgSerProSerI 232
OY 682 -----TCAT 685
Db 232 IeProHisGlnTyrlTyrlProLeuThrLeuGlnTyrlLeuLeuLysHisPhePheLysLeuSer 252
OY 686 TCACGACAGACAGACACAGCTGTGCAGAAACAGATCATAGAGAGAGCAGCCTGTG 745
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 252 InthSerSerLysAsnLeuAlaAlaArg-ValLeuSerGluIlePheSerProMet 271
OY 746 CTCCTCAGA-----GATCCAGCCTCCAGCCAGCTGACGCTCAGCAA---ATATAC 796
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 272 LeupheArgPheSerAlaAlaSerSerAspAsnThrGluAsnLeuIleLysValIleGlu 291
OY 797 TGCCCTATTAATACACACACCTGAGAG-----CCACCTTTCC 832
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 292 IleLeuIleSerThrGlu--TrpAsnGluArgGlnProAlaProAlaLeuProPolysP 311
OY 833 CCAAAAG-----TCTATGAAAGAGATGTCACACACA 865
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 311 roProLysProThrValAlaAsnAsnGlyMetAsnAsnAsnMetSerLeuGlnAsnA 331
OY 866 ATGAATGTCACATGTGAGAAATACAGCCGAGAGAGTGAAGAGCATTTCATGAAGAGAGA 925
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 331 IagLIntrPtyrTrpGlyAspIleSerArgGluGluValAsnGluLys---LeuArgAspT 350
OY 926 ACAAGATGTAGTTCTTGTCGCGATGTTCACAAATCCAGAGAGAGCCCTATG 985
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 350 hrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp---TyT 369
OY 986 TTTTGGCTGTGTTTATGAGAACAAAGTCTACATGTAAA---ATCCGCTTCTTGAGA 1042
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 369 hrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIlePheHisA 386
OY 1043 GGAATGACAGATTTGGCTGGGAGAGAGACTCAGAGAGATGAGAGTTGATTGATCAGT 1102
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 386 rgsAspGlyLysTyrlGlyPheSerAspProLeu-----ThrPheSerValVal 402
OY 1103 AAGACATCATCGAACACTACAAAGAT 1128
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 402 alGluLeuIleAsnHisTyrlArgAsn 410

```

RESULT 5
 US-08-539-005-2
 ; Sequence 2, Application US/08539005
 ; Patent No. 5858686

```

GENERAL INFORMATION:
APPLICANT: Schllessinger, Joseph
APPLICANT: Skolnick, Edward Y.
APPLICANT: Margolis, Benjamin L.
TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
NUMBER OF INVENTION: KINASES AND NOVEL TARGET PROTEINS
CORRESPONDENCE ADDRESS:
ADDRESS: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: 10036-2711
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,005
FILING DATE: 4-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/167,035
FILING DATE: 16-DEC-1993
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-539-005-2
Alignment Scores:
Pred. No.: 1,79e-07 Length: 724
Score: 155.00 Matches: 94
Percent Similarity: 37.24% Conservative: 52
Best Local Similarity: 23.98% Mismatches: 103
Query Match: 7.67% Indels: 143
Gaps: 20
US-09-856-061-3 (1-1129) x US-08-539-005-2 (1-724)
OY 308 GACAGCCGACCTGGAACACACAGACGAGTTGGAAGAGTGCAGAAC----- 355
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 47 GluAlaArgPro-----ArgArgAsnGlyTrpLeuAsnGlyTyrlAsnGlnThrGly 64
OY 356 -----CCATTCGAGGAGACGTGAGAACCAAAACATTAAAGAGATGATCCGTA 406
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 65 GluLysGlyAspPheProGlyThr-----TyValGlnTyrlleGly-Argly 80
OY 407 GAAGAACAAAGATTCCTTACACACCTCCTGCGCTTCATTAACACTTCGAGAGATAC 466
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 80 slsIleSerPro-ProThrProLysProArgPro-----ProArg----- 93
OY 467 AACCTTGCCTTGAGCCGAGAGACAGCAGG----- 498
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 94 --ProleuproValaIalaproGlySerSerlysthrGluAlaaspValglungInGlnAlaI 113
OY 499 -----CCACCTTTATCTCAGACACACACCTTTCCAGAAATC----- 534
    |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 113 euthrleuproaspIleuAlaglInGlnpheAlaproProaspIleAlaproProleuLeuI 133

```

```

QY 535 -----CAGGGAATG-----C 544
DB 133 lclysleuvalglualalegllulyslglyleuclucysserthrleutyargthg 153
QY 545 CCAGTCAGATAGCTTAAGGAGCTTAAGTGAAGCTCTTGAACGAGAAAAGTTCCAT 604
DB 153 lnserserslnleuallaglluleuarglileuileuaspysasp---ThrProser 172
QY 605 ACCAAGAGAGCCCTGAATCACTCATCTGTTAGAAAACCA----- 645
DB 112 alaspheuglmetlleasprvalhisvalleuallaaspralaphelusargtyrleu 192
QY 646 -----AATACCTCAAGATTCACCTTGCCATTAGCAGTTCT----- 681
DB 192 splleuproasnprovalleproallaalaivaltyrserglumetlleserleuala 212
QY 681 ----- 681
DB 212 luvalglnsersercluclytyrileglnleuileuylsleuileargerproser 232
QY 682 -----TCAT 685
DB 232 leronhslglnttytrpleuthrleuuglnlyrleuileuylshispherhelysleu 252
QY 686 TCACGACAGACAGACAGCTGTCMAAACAGAGATCATAGAGAGGATGACCCCTGT 745
DB 252 lnthrserlyasnleuileuasnalaarg-valleuserglulilepherproket 271
QY 746 CTCCTCAAA-----GATGCCAGCTCCAGCCAGCTGCAGCCCTCAGCAA--ATATAC 796
DB 272 leupheargphersealalaalaserseaspranthrgluaaleuilelyvalilegl 291
QY 797 TGCCCTATAATACAGAGCTGAGAA-----CCACCTTCC 832
DB 292 lleuileleserthnglu--trpansgluarglinproallaaleuileuileu 311
QY 833 CCAAAAGG------TCTGATAGAAAGAGATGTCAGAC 865
DB 311 ropolysprothrthryalalaasnanglymetlaasnasnmetserleuglnasna 331
QY 866 ATGATGCTACATGAGCAATACAGCCGCGAGCGAGAGAGAGATTCATGAGAGGA 925
DB 331 laglutrptytrleuylasprleserarglucgluvalasnclulys--leuargasp 350
QY 926 ACAAGATGAGTGTCTTGTCGAGATGTCACAAATCCAGAGAGAGCCCTATG 985
DB 350 hrallaaarglythrphelvalalargaspralaserthrlysmethisgllyasp---tyr 369
QY 986 TTTGGCTGTGTTTATGAGAACAAAGCTTCATGTAA--ATCCGCTCTCGAGA 1042
DB 369 hrleuthrleu-----arglysglyllyasnasnlyleuileylslepherhisa 386
QY 1043 GGAATCAGAGATTTGCGCTGGGAGACAGACTCAGAGAGATGAGAGATGATTCATG 1102
DB 386 rgasprgllystyrglyrphesersproleu-----Thrpheserservalv 402
QY 1103 AAGCATCATGCAACTACAGAAAT 1128
DB 402 alglileuileasnhislytargasn 410

```

RESULT 6
US-09-280-598-5

Sequence 5, Application US/09280598
Patent No. 6391584
GENERAL INFORMATION:
APPLICANT: Schlessler, Joseph
APPLICANT: Skolnik, Edward Y.
APPLICANT: Margolis, Benjamin L.
APPLICANT: App. Harold
TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
KINASES AND NOVEL TARGET PROTEINS

```

NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/280,598
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/252,820
FILING DATE: 02-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-067
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-280-598-5

Alignment Scores:
Score: 5.24e-07 Length: 724
Percent Similarity: 150.50 Matches: 82
Best Local Similarity: 37.048 Conservative: 48
Query Match: 23.368 Mismatches: 93
DB: 7.458 Indels: 128
Gaps: 17

US-09-856-061-3 (1-1129) x US-09-280-598-5 (1-724)
QY 406 AGAAGAACAAGATTCCTTACCACTTCGCGCTCATACACTTCCGAGAGATAC 465
DB 80 lylslylleserprothrprolysprolargpro-----Proarg----- 93
QY 466 CAACCCCTTGCCCTTGAGCCGAGAGAGCAGG----- 498
DB 94 ---ProleuProvalalaprogllyserlystrhualaaaspralaglucglnala 112
QY 499 -----CCACCTTATCTCAGAGACACACCTTTCAGAGATC----- 534
DB 113 leuthrleuproasprleualaglucglnphealaproproasprleualaproproleu 132
QY 535 -----CAGGGAATG----- 543
DB 133 lclysleuvalglualalegllulyslglyleuclucysserthrleutyargthr 152
QY 544 CCAGTCAGATAGCTTAAGGAGCTTAAGTGAAGCTCTTGAACGAGAAAAGTTCCAT 603
DB 153 glnserserslnleuallaglluleuarglileuileuaspysasp---ThrProser 171
QY 604 ACCAAGAGAGCCCTGAATCACTCATCTGTTAGAAAACCA----- 645
DB 172 valaspheuglmetlleasprvalhisvalleuallaaspralaphelusargtyrleu 191
QY 646 -----AATACCTCAAGATTCACCTTGCCATTAGCAGTTCT----- 681
DB 192 splleuproasnprovalleproallaalaivaltyrserglumetlleserleuala 211

```

```

QY 681 ----- 681
Db 212 GluValGlnSerSerGluGluTyrIleGlnLeuLeuLysLysLeuIleArgSerProSer 231
QY 682 -----TCA 684
Db 232 IleProHisGlnTyrTrpLeuThrLeuGlnTyrLeuLeuLysHisPhePheLysLeuSer 251
QY 685 TTCACGACAGACACACAGTGTGCAAAACAGATCATAGAGAGGCATGCACCTTGT 744
Db 252 GlnThrSerSerLysAsnLeuLeuAsnAlaArg-ValLeuSerGluIlePheSerProSe 271
QY 745 TCTCCACACA-----GATGCCAGCTCCAGCCAGCTGACGCCCTCAGCAAA--ATATA 795
Db 271 LeuPheArgPheSerAlaIleSerSerAspAsnThrGluAsnLeuIleLysValIleG1 291
QY 796 CTGCCCTATAAATACACAGCTGAGACA-----CCACCTTTC 831
Db 291 uIleLeuIleSerThrGlu--TrpAsnGluArgGlnProAlaProAlaLeuProPolys 310
QY 832 CCCAAAAG-----TCTGATAGAAAAGATGTCCAGCAC 864
Db 311 ProProLysProThrThrValAlaAsnAsnGlyMetAsnAsnMetSerLeuGlnAsn 330
QY 865 AATGATGTGACATGTGAGAAATACAGCCCGCAGCAGTGAAGAGGCATTCATGAAGAG 924
Db 331 AlaGluTrpTyrTrpIleAspIleSerArgGluGluValAsnGluLys--LeuArgAsp 349
QY 925 AACAGAGATGTAGTTCCTGTCGCGAGATGTTCCACAAAATCCAAAGAACGCCCTAT 984
Db 350 ThrAlaAspIleThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp---Tyr 368
QY 985 GTTTGGCTGTGTTTATAGAAACAAGTCTACAATGTAAA--ATCCGCTTCCTGGAG 1041
Db 369 ThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIlePheHis 385
QY 1042 AGGAATCAGCAGTTGGCTGCGGAGACAGCAGTCAAGAGATGAGAAATGATTCAGTA 1101
Db 386 ArgAspIleLysTyrGlyPheSerAspProLeu-----ThrPheSerVal 401
QY 1102 GAAGACATCATCGACACTCAAGAAAT 1128
Db 402 ValGluLeuIleAsnHisTyrArgAsn 410

RESULT 7
US-08-390-874C-12
; Sequence 12, Application US/08390874C
; Patent No. 6043062
; GENERAL INFORMATION:
; APPLICANT: Kimpel, Anke
; TITLE OF INVENTION: A Constitutively Active
; TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/390,874C
; FILING DATE: 17-FEB-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.

```

```

; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 02307K-057000US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300.
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-390-874C-12

Alignment Scores:
Pred. No.: 2,19e-06 Length: 722
Score: 144.50 Matches: 75
Percent Similarity: 32.76% Conserved: 40
Best Local Similarity: 21.37% Mismatches: 95
Query Match: 7.15% Indels: 141
DB: Gaps: 17

US-09-856-061-3 (1-1129) x US-08-390-874C-12 (1-722)
QY 421 CCTTACACACCTCTCGGCTCTCATACACTTCGAGAGATGCCAGTATACCTTGCCTT 480
Db 85 ProThrProLysProArgPro-----ProArg-----ProLeuProVal 97
QY 481 GAGCCGAGAGCAGCAGG-----CCACCT 504
Db 98 AlaProGlySerSerLysThrGluAlaAspThrGluGlnAlaLeuProLeuProAsp 117
QY 505 TTATCTCAGACACACACCTTCCAGAAATGCCAGGAATGCCAGTATACCTTAAAG 564
Db 118 LeuAlaGluGlnPheAlaProProAspVal---AlaProProLeuLeuIleLysLeu 136
QY 565 GACTTAAGTAGGTC-----CTTGAAGCAGAAAAGTTCCTCATACACAGCAAGCCT 618
Db 137 GluAlaIleGluLysLysGlyLeuGluLysSerThrLeuThrArgThrGlnSerSer 156
QY 619 GAATCAACTATCTGTAGAA-----AACCAAAATACTCAAGAG 657
Db 157 AsnProAlaGluLeuArgGlnLeuLeuAspCysAspAlaAlaSerValaAspLeuGluMet 176
QY 658 ATTCCACTGCCATTTACAGCTTCTCATTCACGACAGCAACACACAGCTGCAAAACAGA 717
Db 177 IleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuAlaAspLeuProAsn 195
QY 718 GATCATAGAGAGGAGCATGACGCCCTGTCCTC----- 750
Db 196 -----ProValIleProValAlaValaTyrAsnGluMetSer 208
QY 751 -----CAGAGATGCCAGCCTCCAGCCAGCTGC----- 777
Db 209 LeuAlaGlnGluLeuGlnSerProGluAspCysIleGlnLeuLeuLysLysLeuIleArg 228
QY 778 -----AGCCCTCAGAAAATATATAGTCCCTATATAATAC----- 810
Db 229 LeuProAsnIleProHisGlnGlyStrpLeuThrLeuGlnTyrLeuLeuLysHisPhePhe 248
QY 810 ----- 810
Db 249 LysLeuSerGlnAlaSerSerLysAsnLeuLeuAsnAlaArgValLeuSerGluIlePhe 268
QY 810 ----- 810
Db 269 SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrGluHisLeuIleLys 288
QY 811 -----ACAAGCTGG----- 819
Db 289 AlaIleGluLeuLeuIleSerThrGluTrpAsnGluArgGlnProAlaProAlaLeuPro 308
QY 820 ---AGACCACCTTCCCC-----AAAAGCTGTAGTAAAGATGTC 858

```

```

Db      309  ProLysProProLysProThrThrValAlaAsnAsnSerMetAsnAsnMetSerLeu 328
           :::::  :::::  :::::
QY      859  CAGCACAATGATGTGATTCATGAGAAATACAGCCCGCAGCAGTGAAGAAGCATTCATG 918
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      329  GlnAspAlaGluTrpTrpTrpGlyAspIleSerArgGluValAsnGluLys-----Leu 346
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      919  AAGGAGACAAGAGATGATGTTCTGTGTCGAGATGTTCCACAAATCCAAAGAAAGAG 978
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db      347  ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrIlysmethHisGlyAsp 366
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      979  CCTATGTTTGGCTGTGTTTATGAGAACAAAGTCTACATGTAA--ATCCGCTTC 1035
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      367  ---TyrThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLys 382
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      1036  CTGAGAGAAATCAGCATGTTGCCCTGGGACAGCAGCTCAGAGAGATGAGAACTTGAT 1095
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      383  PheHisArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheAsn 398
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY      1096  TCAGTAGAAGACATCATCGAACATCAAGAAAT 1128
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      399  SerValValGluLeuIleAsnHisTyrArgAsn 409
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```

RESULT 8

```

US-09-265-772-12
; Sequence 12, Application US/09265772
; Patent No. 6300111
; GENERAL INFORMATION:
; APPLICANT: Klippel, Anke
; APPLICANT: Williams, Lewis T.
; TITLE OF INVENTION: A Constitutively Active
; TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,772
; FILING DATE: 10-MAR-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,874
; FILING DATE: 17-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 02307K-057010US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-265-772-12

```

```

Alignment Scores: 2.19e-06 Length: 722
Score: 144.50 Matches: 75
Percent Similarity: 32.76% Conservative: 40
Best Local Similarity: 21.37% Mismatches: 95

```

```

Query Match: 7.15% Indels: 141
DB: 4 Gaps: 17
US-09-856-061-3 (1-1129) x US-09-265-772-12 (1-722)
QY 421 CCTTACACACGTCGCGGCTCTCATTAACATTCGGAAGAACTACCAACCTTGGCCCT 480
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 85 ProThrProLysProArgPro-----ProArg-----ProLeuProVal 97
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 481 GAGCCGAGAGACGAGC-----CTGAGACAGAAAAAGTTCCTCATTAACACAGAGAAAGCT 504
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 98 AlaProGlySerSerLysThrGluAlaAspThrGluGlnAlaLeuProLeuProAsp 117
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 505 TTAATCTCAGACACACACCTTCCAGATCCAGGAATCCAGTCAGATTAAGTTAAAG 564
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 118 LeuAlaGluGlnPheAlaProProAspVal---AlaProProLeuLeuIleLysLeuLeu 136
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 565 GACTTAAGTGAGTC-----CTGAGACAGAAAAAGTTCCTCATTAACACAGAGAAAGCT 618
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 137 GluAlaIleGluLysLysGlyLeuGlyLysSerThrLeuTyrArgThrGlnSerSer 156
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 619 GAATCACTCATCTGTAGAA-----AACCAAAATACTCAAGAG 657
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 157 AsnProAlaGluLeuArgGlnLeuLeuAspCysAspAlaAlaSerValAspLeuGlnMet 176
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 658 ATTCCACTTGGCCATTAAGCATTCCTTCAACAGACAGACACCAACCAAGTCGCAAAACAGA 717
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 177 IleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuAlaAspLeuProAsn--- 195
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 718 GATCATAGAGAGAGCATCAGCCCTGTCTCT 750
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 196 -----ProValIleProValAlaValItyrAsnGlnMetMetSer 208
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 751 -----CAGAGATGCCACCTCCAGCCAGCTGC----- 777
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 209 LeuAlaGlnGluLeuGlnSerProGlnAspCysIleGlnLeuLysLysLeuIleArg 228
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 778 -----AGCCCTCAGCAAAATATAGCCCTATAAATAC----- 810
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 229 LeuProAsnIleProHisGlnCysTrpLeuThrLeuGlnTyrLeuLysHisPhePhe 248
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 810 ----- 810
Db 249 LysLeuSerGlnAlaSerSerLysAsnLeuLeuAsnAlaArgValLeuSerGluIlePhe 268
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 810 ----- 810
Db 269 SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrGlnHisLeuIleLys 288
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 811 -----ACAAGCTGG----- 819
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 289 AlaIleGluIleLeuIleSerThrGluTrpAsnGluArgGlnProAlaProAlaLeuPro 308
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 820 ---AGACCACCTTCCCC-----AAAAGTGTGATGAGAAAGATGTC 858
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 309 ProLysProProLysProThrThrValAlaAsnAsnSerMetAsnAsnMetSerLeu 328
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 859 CAGCACAATGATGTGATTCATGAGAAATACAGCCCGCAGCAGTGAAGAAGCATTCATG 918
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 329 GlnAspAlaGluTrpTrpTrpGlyAspIleSerArgGluValAsnGluLys-----Leu 346
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 919 AAGGAGACAAGAGATGATGTTCTGTGTCGAGATGTTCCACAAATCCAAAGAAAGAG 978
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
Db 347 ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrIlysmethHisGlyAsp 366
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 979 CCTATGTTTGGCTGTGTTTATGAGAACAAAGTCTACATGTAA--ATCCGCTTC 1035
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 367 ---TyrThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLys 382
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::
QY 1036  CTGAGAGAAATCAGCATGTTGCCCTGGGACAGCAGCTCAGAGAGATGAGAACTTGAT 1095
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db 383  PheHisArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheAsn 398
           :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::  :::::

```


QY 1096 TCAGTAGAGACATCATGACACACATCAAGAAAT 1128
 Db 399 SerValValGluLeuIleAsnHisTyrArgAsn 409

RESULT 9

US-08-391-615-2

Sequence 2, Application US/08391615
 Patent No. 550054

GENERAL INFORMATION:

APPLICANT: Witte, Owen

APPLICANT: Tsukada, Satoshi

APPLICANT: Safran, Douglas

APPLICANT: Rawlings, David

TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE

TITLE OF INVENTION:

NUMBER OF SEQUENCES: 7

CORRESPONDENCE ADDRESS:

ADDRESSEE: FLEHR, HOHBACH, TEST, ALBRITTON & HERBERT

STREET: 4 Embarcadero Center, Suite 3400

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-4187

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/391,615

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/006,449

FILING DATE: 21-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Rowland, Berttram I

REGISTRATION NUMBER: 20,015

REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 781-1989

TELEFAX: (415) 398-3249

TELEX: 910 277299 FHT UR

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 659 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-391-615-2

Alignment Scores:

Pred. No.: 2,996-06

Score: 143.00

Percent Similarity: 33.98%

Best Local Similarity: 23.44%

Query Match: 7.08%

Db: 1

Gaps: 11

US-09-856-061-3 (1-1129) x US-08-391-615-2 (1-659)

QY 373 AGAAGCCAAACATTAAGAGATGATCGATCGTAAGAAGAACAGATTCTTACACACT 432

Db 171 ArgAsnGlySerLeuLysProGlySerSerHisArgLysThrLysLysProLeuPro 190

QY 433 CCTGGCCCTCTCATTAACACTCCGAGAGAGTACCAACCTGGCCCTTGAGCGGAGAGC 492

Db 191 ThrProGluGluAspGlnIleLeuLysLys-----ProLeuProProGluProHla 208

QY 493 AGAGGCCACCTTATCTGAGACACACTTCCAGAGTCCAGGGAATGCCAGTCAG 552

Db 209 -----AlaPro 210

QY 553 ATAACCTTAAGGACCTTAAGTGAGGTCTCTGAGACA---GAAAAGTTCCTCATACAGC 609
 Db 211 IleSerThrThrGluLeuLysValValAlaLeuTyrAspTyrMetProMetAsnIla 230

QY 610 -----AGNAGCCCTGAATCACTCACTCTGTGTGAACCAAAATCACTCA 654
 Db 231 AsnAspLeuGlnLeuArgLysGlyGluLysThrPheIleLeuGluLysSerAsn----- 248

QY 655 GAGATTCACCTTGCATTAAGCAGTCTTCTCATTCAGACACAGACACAGTGTGCAAAAC 714
 Db 249 ---LeuProTrp-----TrpArgAla 254

QY 715 AGAGATCATATAGAGAGCAGCAGCCCTGCTCTCTCAGAGATGCCAGCTTCACCGAC 774
 Db 255 ArgAspLysAsnGly----- 259

QY 775 TGCAGCCCTCAGCAAAATCACTGCTCTTAATATACACAGCTGAGACACCTTCCCC 834
 Db 260 -----GlnGluGlyTyrIleProSerAsnTyrIleThr----- 270

QY 835 AAAAGTCTGATAGAAAGATGTCGACACACATGATGATGATGATGATGATGATGATGAT 894
 Db 271 -----GluAlaGluAspSerIleGluMetTyrGluIlePtyrSerLysHisMetThrArg 288

QY 895 CAGCAGTGGAGAGGACATTCATGAAGAGAACAGAGTGTGCTGCTCGAGAT 954
 Db 289 SerGlnAlaGluGlnLeuLeuLysGlnGluGlyGlyGlnGlyGlyPheIleValArgAsp 308

QY 955 TGTTCACAAATTCACGAGAGACCCCTATGTGTTGGCTGTGTTTATAGAACAAAGTC 1014
 Db 309 SerSerLysAlaGlyLys----- 314

QY 1015 TACATGTAAATATCGCTCTCGAGAGAGATACAGACATTTGCCCTGGAGACAGACTC 1074
 Db 315 TyrThrValSerVal-----PheAlaLysSerThrGly---- 325

QY 1075 AGAGAGATGAGAAAGTTGATTCAGTAGAGACATCATCAACACTAC 1122
 Db 326 -----GluProGlnGlyValIleArgHisTyr 334

RESULT 10

US-08-426-509A-8

Sequence 8, Application US/08426509A

Patent No. 6326469

GENERAL INFORMATION:

APPLICANT: Ullrich, Axel

APPLICANT: Gishizky, Mikhail

APPLICANT: Sures, Itman G.

TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN

TITLE OF INVENTION: TYROSINE KINASES

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York,

STATE: NY

COUNTRY: USA

ZIP: 10036-2711

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/426,509A

FILING DATE: 21-APR-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/232,545

ATTORNEY/AGENT INFORMATION:

NAME: Coruzzi, Laura A

REGISTRATION NUMBER: 30,742

REFERENCE/DOCKET NUMBER: 7683-0074-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: No. 6326469e
US-08-426-509A-8

Alignment Scores:
Pred. No.: 6,12e-06
Score: 140.00
Percent Similarity: 33.98%
Best Local Similarity: 22.66%
Query Match: 6.93%
DB: 4
Matches: 58
Conservative: 29
Mismatch: 71
Indels: 98
Gaps: 11

US-09-856-061-3 (1-1129) x US-08-426-509A-8 (1-659)

QY 373 AGAAGCCAAACATTAAAGGAGATGATCCGTAAGAAAGAACAGATTCTTACCACT 432
DB 171 ArgasnlyserleuylsproglyserhisargylsThrlylserProleupPro 190
QY 433 CCTGGCCTCTCATACACTCCGAGAGATGACCACTTCCGCGCCGAGCCGAGAGC 492
DB 191 ThrProgluLualspglnlleleuyls-----ProleupProgluProAla 208
QY 493 AGAGGCCACCTTATCTCAGAGACACCTTCCAGAGTCCAGGAATGCCAGTCAG 552
DB 209 -----AlaPro 210
QY 553 ATAAGCTTAAGGACTTAAGTGAAGTCTTGAAACA---GAAAAAGTCTCTATACAG 609
DB 211 ValSerThrsergluLeuylsValAlaLeuTyrsPtyrMetPromelaSna 230
QY 610 -----AGGAAGCCTGAATCACTCATCTGTAGAAACCAAAATACTCA 654
DB 231 AsnAspLeuInleuArglysglyAspGlyThrleuLeuInleuSerAsn----- 248
QY 655 GAATTCACACTTGCCATTTACAGCTTCTTCATCAGACAGCAACCAAGTGTCAAAAC 714
DB 249 ---LeupProTrp-----TrpArgAla 254
QY 715 AGAGATCATAGAGGAGCATGACCCCTGTTCTCAGAGATGCCAGCTCCAGCCAGC 774
DB 255 ArgAspLysasnly----- 259
QY 775 TGCAGCCCTCAGAAATATATACTGCCCTATTAATACACAGCTGAGACACCTTCC 834
DB 260 -----GlnGluGlyTyrlleProserAsnTyrlValThr----- 270
QY 835 AAAAGCTGATGAAGAGATGTCACACAAATGAATGATGAGATGAGAAATACAGCCG 894
DB 271 -----GluAlaGluAspserIleGluMetTyrgluTrpTyrserlyshsmethrArg 288
QY 895 CAGGAGCAGAGGAGCATTCATGAGAGAACAGAGATGATGTTCTTGTCGCCAGAT 954
DB 289 serGlnAlaGluInleuLeuylsGlnGlyLysGlnGlyGlyheleValAlaArgasp 308
QY 955 TGTTCACAAAAATCCAGAGAGAGCCCTATGTTTGGCTGTGTTATAGAACAAAGTC 1014
DB 309 SerSerLysAlaGlyLys----- 314
QY 1015 TACAAATGTAAATCCGCTTCTCGAGAGAGATCAGACATTTGCCCTGGGAGACAGACTC 1074
DB 315 TyrlThrValSerVal-----PheAlaLysSerThrGly----- 325
QY 1075 AGAGGAGATGAGAGATTTGATTCAGTAGAAGACATCATGACACACTAC 1122

DB 326 -----AspProGlnGlyValIleArgHisTyr 334

RESULT 11
PCT-US95-05008-8
Sequence 8, Application PC/TUS9505008
GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissenschaften E.V.
APPLICANT: Hofgarten Str. 2
APPLICANT: Munchen 80539
TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
NUMBER OF INVENTION: Kinases
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US95/05008
FILING DATE: 24-Apr-1995
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/232,545
FILING DATE: 22-Apr-1994
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-074
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212)790-9090
TELEFAX: (212)869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 659 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
PCT-US95-05008-8

Alignment Scores:
Pred. No.: 6,12e-06
Score: 140.00
Percent Similarity: 33.98%
Best Local Similarity: 22.66%
Query Match: 6.93%
DB: 5
Matches: 59
Conservative: 29
Mismatch: 71
Indels: 98
Gaps: 11

US-09-856-061-3 (1-1129) x PCT-US95-05008-8 (1-659)

QY 373 AGAAGCCAAACATTAAAGGAGATGATCCGTAAGAAAGAACAGATTCTTACCACT 432
DB 171 ArgasnlyserleuylsproglyserhisargylsThrlylserProleupPro 190
QY 433 CCTGGCCTCTCATACACTCCGAGAGATGACCACTTCCGCGCCGAGCCGAGAGC 492
DB 191 ThrProgluLualspglnlleleuyls-----ProleupProgluProAla 208

```

OY 493 AGCAGGCCACCTTTATCTCAGAGACACACTTTCAGAAAGTCCAGGAATGCCAGTGC 552
DB 209 -----Alapro 210
OY 553 ATAGCTTAAGGACTTAAGTAGGCTCCTTGAAGA---GAAAAAGTCCATACACAG 609
DB 211 ValSerThrSerGluLeuLysValAlaLeuLysPheMetProMetAla 230
OY 610 -----AGAAAGCTGAATCACTCATCTCTGTGAAGAAACCAATACTCA 654
DB 231 AsnAspLeuGlnLeuArgLysGlyAspLysPheLeuGlnLysSerAsn----- 248
OY 655 GAGATTCACCTGTCATAGCAGTTCCTTCATTCAGACAGCAACACACAGTGTGCAAAAC 714
DB 249 ---LeuProTrp-----TrpArgAla 254
OY 715 AGAGATCATAGAGAGCATGCAGCCCTGCTCTCCTCAGAGATGCCAGCTCCAGCCAGC 774
DB 255 ArgAspLysAsnGly----- 259
OY 775 TGCACCCCTCAGCAAAATATATCTGCCCTTAATAATACACAGCTGGAGACACCTTCCCC 834
DB 260 -----GlnGlnGlyTyrLysProSerAsnTyrValThr----- 270
OY 835 AAAAGTCTGATAGAAAGGATGTCACGACCAATGAATGTGATGATGGAATAACAGCCGC 894
DB 271 -----GluAlaGlnAspSerLysGlnMetTyrGlnTyrPheLysHisMetThrArg 288
OY 895 CAGGACGTGAGAGAGCATTCATGAAGAGAAACAGAGATGATGTTCTGTGTCGAGAT 954
DB 289 SerGlnAlaGlnGlnLeuLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 308
OY 955 TCTTCACCAAAATCCAGAGAAAGCCCTATGTTTGGCTGTGTTTATAGAAACAAAGTC 1014
DB 309 SerSerLysAlaGlyLys----- 314
OY 1015 TACAAATGTAAATCCGCTTCTCTCGAGAGAAATCAGACAGTTTGCCTGGAGACAGCAGTC 1074
DB 315 TyrThrValSerVal-----PheAlaLysSerThrGly----- 325
OY 1075 AAGAGGATGAGAGATGTTGATTCAGTAGAAGACATCATGACACACTAC 1122
DB 326 -----AspProGlnGlyValIleArgHisTyr 334

RESULT 12
US-08-391-615-6
; Sequence 6, Application US/08391615
; Patent No. 5550054
; GENERAL INFORMATION:
; APPLICANT: Witte, Owen
; APPLICANT: Tsukada, Satoshi
; APPLICANT: Saitan, Douglas
; APPLICANT: Rawlings, David
; TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOBBACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,615
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:

```

```

APPLICATION NUMBER: US 08/006,449
FILING DATE: 21-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Rowland, Bertram I
REGISTRATION NUMBER: 20,015
REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 781-1989
TELEFAX: (415) 398-3249
TELEX: 910 277299 FHT UR
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 442 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-391-615-6

Alignment Scores:
Pred. No.: 1,036-05 Length: 442
Score: 137.00 Matches: 60
Percent Similarity: 34.50% Conservative: 29
Best Local Similarity: 23.26% Mismatches: 67
Query Match: 6.78% Indels: 102
DB: 1 Gaps: 11

US-09-856-061-3 (1-1129) x US-08-391-615-6 (1-442)
OY 373 AGAAGCCAAAGATTAAAGAGATGATCGTATCGTAAGAAAGACAGATTCCTTACCACT 432
DB 171 ArgAsnGlySerLeuLysProGlnSerSerHisArgLysThrLysLysProLeuPro 190
OY 433 CCTGGCCTCTCATTAACACTTCCGAAGAAGTACCAACCTTGCCCTGAGCCGGAGAGC 492
DB 191 ThrProGlnLysAspGlnLeuLysLys-----ProLeuProGlnProAlaAla 208
OY 493 AGCAGGCCACCTTATCTCAGACACACCTTCCAGAGTC-----CAGGAGAT 543
DB 209 Ala-----ProValSerThrSerGlnLeuLysLysGlnValAlaLeuLysPheLysMet 226
OY 544 CCC-----AGTCAGATTAAGCTTAAGGACTTAAGTGAG-----GTCTTGAAGCA 588
DB 227 ProMetAsnAlaAsnAspLeuGlnLeuArgLysGlyAspLysPheLeuGlnGln 246
OY 589 GAAAAAGTCTCTCATTAACAGAGAGAGCCGTGAATCAATCATCTGTTAGAAACCAAAAT 648
DB 247 SerAsnLeuProThrPheArgAlaArg-----AspLysAsn 258
OY 649 ACTCAGAGATTCACCTTGCATTAGCAGTTCCTTCATTCAGACAGCAACACAGCTGTG 708
DB 259 GlyGlnGln-----GlyTyrLysProSerAsnAspValThr 270
OY 709 CAAACAGAGATCATAGAGAGAGAGCAGCCCTGTCTCTCAGAGATGCCAGCTCCA 768
DB 271 GluAlaGlnAsp----- 274
OY 769 GCCAGCTGCAGCCCTCAGCAAAATATACCTGCTTATTAATACAAAGCTGGAGACCACT 828
DB 274 ----- 274
OY 829 TTCCCCAAAGTCTGATAGAAAGATGCCAGACATGATGATGATGGAATAC 888
DB 275 -----SerLysGlnMetTyrLysPheLysMet 286
OY 889 AGCCGCGAGCAGCTGGAAGAGCATTCATGAAGAGAAACAGATGATGATGCTGTGCTC 948
DB 287 ThrArgSerGlnAlaGlnGlnLeuLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 306
OY 949 CGAGATTTGTTCCACAAATCCAAAGAGAGAGCCCTATGTTTGGCTGTGTTTATGAGAAC 1008
DB 307 ArgAspSerSerLysAlaAlaLys-----TyrThrLeuSerVal----- 319

```



```
OY 853 GATGTCACACATGATGTACATTGGAGATACGCCGACGACAGAGAGCA 912
      :::::||||| ||||| ||||| ||||| |||||
Db 493 GluLeuGlnAlaGluThrTrpTyrGlnGlyGluMetSerArgLysGlnAlaGluGlyLeu 512
      :::::||||| ||||| ||||| ||||| |||||
OY 913 TTCATGAGAGAGACACAGCATGCTACTTCTTGTCGAGATTGTTCCACAAAATCCAG 972
      ||||| ||||| ||||| ||||| |||||
Db 513 Leu-----GluLysAspGlyAspPheLeuValArg-----LysSerThr 525
      ||||| ||||| ||||| ||||| |||||
OY 973 GAAGAGCCC-----TATGTTTGGCTGCTTTTATGAGACAAAGCTACATGTAAAA 1026
      ||| ::::: ||||| ||||| ||||| |||||
Db 526 ThrAsnProGlySerPheValLeuThrGlyMetHisAsnGlyGlnAlaLysHisLeu 545
      ::::: ||||| ||||| ||||| ||||| |||||
OY 1027 ATCCGCTTCCTGGAGAGCAATCAGCAGTTGCCCTGGGGACAGACTCAGAGGAGATGAG 1086
      ::::: ||||| ||||| ||||| ||||| |||||
Db 546 LeuValAspProGlu-----GlyThrIleArgThrLysAspArg 558
      ||||| ||||| ||||| ||||| |||||
OY 1087 AAGTTGATTGACTAGACATCATCCAGACACTAC 1122
      ||||| ||||| ||||| ||||| |||||
Db 559 ValPheAspSerIleSerHisLeuIleAsnHisHis 570
      ||||| ||||| ||||| ||||| |||||
```

Search completed: April 21, 2003, 12:33:09
Job time : 25.2707 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:20:58 ; Search time 28.7202 Seconds

(without alignments)
5944.816 Million cell updates/sec

Title: US-09-856-061-3

Perfect score: 2020
Sequence: 1 ttccagaacttcagctgc.....catcgaaactacagaatt 1129

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 288829 seqs, 7561385 residues

Total number of hits satisfying chosen parameters: 577658

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+.n2p.model -DEV=xlp
-O=/cgn2_1/USPTO.spool/US0985661/runat.21042003_113021_3809/app.query.fasta.1.3150
-DB=Published.Applications_AA -QFMT=fastan -SUFFIX=n2p.rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTPMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US0985661.ecgn.1.1.34_etrunat.21042003_113021_3809
-NCPU=6 -ICPU=3 -NO.XLPHY -NO.WMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/PTUS_PUBCOMB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	2009	99.5	428	US-09-966-955A-4
2	2009	99.5	443	US-09-966-955A-2
3	1694	83.9	353	US-09-966-955A-6
4	238	11.8	46	US-09-864-761-38652

5	214	10.6	40	US-09-864-761-44023	Sequence 44023, A
6	214	10.6	40	US-09-864-761-44024	Sequence 44024, A
7	166	8.2	30	US-09-864-761-38653	Sequence 38653, A
8	150.3	7.5	724	US-09-962-925-4	Sequence 4, Appl1
9	149	7.4	688	US-10-081-980B-1	Sequence 1, Appl1
10	149	7.4	724	US-10-081-980B-3	Sequence 3, Appl1
11	147	7.3	724	US-09-962-929-2	Sequence 2, Appl1
12	144.5	7.2	945	US-08-965-927-2	Sequence 2, Appl1
13	144.5	7.2	945	US-10-121-882-2	Sequence 2, Appl1
14	143	7.1	659	US-10-045-202-4	Sequence 4, Appl1
15	140	6.9	659	US-09-977-260-8	Sequence 8, Appl1
16	140	6.9	659	US-10-045-202-2	Sequence 2, Appl1
17	140	6.9	659	US-09-977-261-8	Sequence 8, Appl1
18	140	6.9	659	US-09-977-269-8	Sequence 8, Appl1
19	126	6.2	597	US-10-038-010-22	Sequence 22, Appl1
20	125.5	6.2	533	US-09-908-805B-63	Sequence 63, Appl1
21	125	6.2	533	US-09-920-021A-3	Sequence 3, Appl1
22	120	5.9	939	US-09-926-248B-29	Sequence 29, Appl1
23	120	5.9	939	US-09-801-368-380	Sequence 380, App
24	119	5.9	620	US-09-977-260-9	Sequence 9, Appl1
25	119	5.9	620	US-09-977-261-9	Sequence 9, Appl1
26	119	5.9	620	US-09-977-269-9	Sequence 9, Appl1
27	115	5.7	675	US-10-186-399-3	Sequence 3, Appl1
28	115	5.7	675	US-09-977-260-4	Sequence 4, Appl1
29	115	5.7	675	US-09-977-261-4	Sequence 4, Appl1
30	115	5.7	675	US-09-977-269-4	Sequence 4, Appl1
31	112.5	5.6	552	US-09-880-192-50	Sequence 50, Appl1
32	112	5.5	1325	US-09-864-761-35612	Sequence 35612, A
33	111	5.5	482	US-09-764-853-568	Sequence 568, App
34	111	5.5	542	US-09-908-805B-44	Sequence 44, Appl1
35	111	5.5	655	US-09-205-658-57	Sequence 57, Appl1
36	111	5.5	655	US-09-844-353A-57	Sequence 57, Appl1
37	111	5.5	944	US-09-964-238-2	Sequence 2, Appl1
38	110	5.4	562	US-09-925-299-869	Sequence 869, App
39	110	5.4	562	US-09-925-299-869	Sequence 869, App
40	110	5.4	1341	US-10-076-622-565	Sequence 565, App
41	110	5.4	1341	US-10-076-622-565	Sequence 565, App
42	110	5.4	1349	US-10-076-622-573	Sequence 573, App
43	110	5.4	1349	US-10-076-622-573	Sequence 573, App
44	109.5	5.4	4019	US-09-854-133-425	Sequence 425, App
45	109.5	5.4	4019	US-09-738-973-425	Sequence 425, App

ALIGNMENTS

RESULT 1
US-09-966-955A-4
Sequence 4, Application US/09966955A
Patent No. US20020155563A1
GENERAL INFORMATION:
APPLICANT: Perez-Villar, Juan J.
APPLICANT: Chang, Han
APPLICANT: Yang, Wen-Pin
APPLICANT: Wu, Yuli
APPLICANT: Whitney, Steven S.
TITLE OF INVENTION: Identification and Cloning of a Full-Length Human
TITLE OF INVENTION: Clnk-related Gene, MIST (Mast Cell Immunoreceptor)
TITLE OF INVENTION: Signal Transducer
FILE REFERENCE: 3053-4113051
CURRENT APPLICATION NUMBER: US/09/966, 955A
CURRENT FILING DATE: 2001-09-28
PRIOR APPLICATION NUMBER: 60/237030
PRIOR FILING DATE: 2000-09-29
NUMBER OF SEQ ID NOS: 52
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 428
TYPE: PRT
ORGANISM: HUMAN
FEATURE:
OTHER INFORMATION: HUMAN MIST SPICE VARIANT CLONE #7, AMINO ACID
OTHER INFORMATION: SEQUENCE

US-09-966-955A-4

Alignment Scores:

Pred. No.:	1,68e-162	Length:	428
Score:	2009.00	Matches:	374
Percent Similarity:	99.73%	Conservative:	1
Best Local Similarity:	99.47%	Mismatches:	1
Query Match:	99.46%	Indels:	0
DB:	9	Gaps:	0

US-09-856-061-3 (1-1129) x US-09-966-955A-4 (1-428)

```

QY 1 TTCGAGAACTTCAGTCTGCCAAAAAAGGTCATGGCGCTGCATCACTAGTCCACAGGC 60
    |||||||
DB 19 PheGlnAnpHeserLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGly 38
QY 61 CAGTACAGAGATGAACAAGCCTCTTCTAGACTGGGAAAGAACTTGGTGCAGTCCG 120
    |||||||
DB 39 GlnTyrGlnArgMetAsnLysProLeuLeuAspTrpGlnArgAsnPheAlaValLeu 58
QY 121 GATGAGCAAAAAGCCACAGCTGATGATGATGATGATGATGATGATGATGATGATG 180
    |||||||
DB 59 AspGlyAlaLysGlyHisSerAspAspArgPtyrAspAspProGluLeuArgMetGluGlu 78
QY 181 ACATGGCAGTCCGATTAATAATTTCACAGCCCGCTATTAAGAAATCTGAATATGAGAT 240
    |||||||
DB 79 ThrTrpGlnSerIleLysIleLeuProAlaArgProIleLysGluSerGluTyrAlaAsp 98
QY 241 ACACACTAATTTCAAGGTTGCAATGACACTCCCTCCGTTGACACCAAGACCTCTATC 300
    |||||||
DB 99 ThrHisTyrPheLysValAlaMetAspTrpProLeuProLeuAspThrTrpTrpSerIle 118
QY 301 TCCATTGGACAGCCGACCTGGAAACACACAGACAGGTTGGAAAGAGTGGCAAAACCAT 360
    |||||||
DB 119 SerIleGlyGlnProThrTrpAsnThrGlnThrArgLeuGlnArgValAspLysProIle 138
QY 361 TCCAGGAGGTCAGAGCCCAAAACATTAAAGAGATGCTCCGTTAAAGAAACAAGATT 420
    |||||||
DB 139 SerLysAspValArgSerIleAsnIleLysGlyAspAlaSerValArgLysAsnLysIle 158
QY 421 CCTTTACACACTTCCTGGGCTCTCATACACTCCGAAAGAACCAACCTTGGCCCT 480
    |||||||
DB 159 ProLeuProProProArgProLeuIleThrProLysLysTyrGlnProLeuProPro 178
QY 481 GAGCCGAGAGACAGACGCCCTTATCTCAGACAGACACACCTTTCCAGAACTCCAGGA 540
    |||||||
DB 179 GluProGluSerSerArgProProLeuSerGlnArgHisThrPheProGluValGlnArg 198
QY 541 ATGCCAGTCAGATTAAGCTTAAGGACTTAAGTGAAGCTCTGAAGCAGAAAAAGTTCT 600
    |||||||
DB 199 MetProSerGlnIleSerLeuArgAspLeuSerGluValLeuGluAlaGluLysValPro 218
QY 601 CATAAACAGAGAGACCTGAATCAACTCATCTGTTAGAAAACCAAAATCTCAGAGAT 660
    |||||||
DB 219 HisAsnGlnArgLysProGluSerThrHisLeuLeuGlnAsnGlnHisThrGlnIle 238
QY 661 CCACTTGGCATTAAGAGTTCTTTCATTCACAGACAAACACAGCTGTGCAAAAAGAT 720
    |||||||
DB 239 ProLysAlaIleSerSerSerSerPheThrTrpSerAsnHisSerValGlnAsnArgAsp 258
QY 721 CATAGAGAGGATGACAGCCCTGTTCTCTCTCAGAGATGACAGCCCTCCAGCCAGCTGC 780
    |||||||
DB 259 HisAlaGlyGlyLeuMetGlnProCysSerProGlnArgCysGlnProProLysSer 278
QY 781 CCTCAGCAAAATATAGTCCCTATTAATAATACACAAGCTGAGAGACACCTTCCCAAAAAG 840
    |||||||
DB 279 ProHisGlnAsnIleLeuProTyrGlyTrpTrpSerTrpArgProProPheProLysArg 298
QY 841 TCTGATAGAAAGATGTCACAGACAATGATGATCATTTGAGAAATACAGCCGACAGCA 900
    |||||||
DB 299 SerAspArgLysAspValGlnHisAsnGlnTyrGlyIleGlyLysSerArgGlnAla 318
QY 901 GTGGAAGAGGATTCATAGAGAGAAACAAGATGCTAGTTCTTGGTCCGAGATTTGTTCC 960
    |||||||

```

```

DB 319 ValGlnGlnAlaPheMetLysGlnAsnLysAspGlySerPheLeuValArgAspCysSer 338
QY 961 ACAAAATCCAGAGAGAGCCCTATGTTTGGCTGTCTTTATGAGAAACAAGTCTACAA 1020
    |||||||
DB 339 ThrLysSerLysGlnGluProTyrValLeuAlaValPheTyrGlnLysnLysValTyrAsn 358
QY 1021 GTAAAAATCCGCTTCTCTGAGAGAGATCAGACAGTTTGGCCCTGGGAGCAGACTCAGAGA 1080
    |||||||
DB 359 ValLysIleArgPheLeuGlnArgAsnGlnGlnPheAlaLeuGlyThrGlyLeuArgGly 378
QY 1081 GATGAGAAATTTGATTCAGTAGAGACATCATGCACTACAGAACTACAAAGAT 1128
    |||||||
DB 379 AspGlnLysPheAspSerValGlnLysPheIleIleGlnHisTyrLysAsn 394

RESULT 2
US-09-966-955A-2
; Sequence 2, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gena S.
; APPLICANT: Kanner, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-length Human
; TITLE OF INVENTION: Clink-related Gene, MIST (Mast Cell Immunoreceptor
; FILE REFERENCE: Signal Transducer)
; CURRENT APPLICATION NUMBER: US/09/966,955A
; PRIOR FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN FULL-LENGTH MIST cDNA CLONE #8, TRANSLATED
; OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-2

Alignment Scores:
Pred. No.: 1.7e-162 Length: 443
Score: 2009.00 Matches: 374
Percent Similarity: 99.73% Conservative: 1
Best Local Similarity: 99.47% Mismatches: 1
Query Match: 99.46% Indels: 0
DB: 9 Gaps: 0

US-09-856-061-3 (1-1129) x US-09-966-955A-2 (1-443)
QY 1 TTCGAGAACTTCAGTCTGCCAAAAAAGGTCATGGCGCTGCATCACTAGTCCACAGGC 60
    |||||||
DB 34 PheGlnAnpHeserLeuProLysAsnArgSerTrpProArgIleAsnSerAlaThrGly 53
QY 61 CAGTACAGAGATGAACAAGCCTCTTCTAGACTGGGAAAGAACTTGGTGCAGTCCG 120
    |||||||
DB 54 GlnTyrGlnArgMetAsnLysProLeuLeuAspTrpGlnArgAsnPheAlaValLeu 73
QY 121 GATGAGCAAAAAGCCACAGCTGATGATGATGATGATGATGATGATGATGATGATG 180
    |||||||
DB 74 AspGlyAlaLysGlyHisSerAspAspArgPtyrAspAspProGluLeuArgMetGluGlu 93
QY 181 ACATGGCAGTCCGATTAATAATTTCACAGCCCGCTATTAAGAAATCTGAATATGAGAT 240
    |||||||
DB 94 ThrTrpGlnSerIleLysIleLeuProAlaArgProIleLysGluSerGluTyrAlaAsp 113
QY 241 ACACACTAATTTCAAGGTTGCAATGACACTCCCTCCGTTGACACCAAGACCTCTATC 300
    |||||||

```



```

Db 114 ThrHisTyLrPhelysValAlaMetAspThrProLeuProLeuAspThrArgThrSerIle 133
QY 301 TCCATGGACAGCCGACCTGGAGACAGAGAGGTGGGAAGAGTGGCAACCACTATP 360
Db 134 SerIleGlyLeuProHisThrPasnThrGlnThrArgLeuGlnArgValAspIle 153
QY 361 TCCAGGAGCTCAGAGCCCAAAACATTAAAGAGATGATCCCTAGAAAGAACAGATT 420
Db 154 SerLysAspValArgSerGlnAsnIleLysGlyAspAlaSerValArgLysAsnLysIle 173
QY 421 CCTTACACACCTCTGGCCCTCTCATPACACTTCGAGAGATGACCAACCTTGGCCCT 480
Db 174 ProLeuProProArgProLeuIleThrLeuProLysIleGlnProLeuProPro 193
QY 481 GACCCGAGAGCAGAGCCGACCTTATCTCAGAGACACACCTTCCAGAGTCCAGGGA 540
Db 194 GluProGlnSerSerArgProProLeuSerGlnArgHisThrPheProGlnValGlnArg 213
QY 541 ATGCCAGTCAGATAAGCTTAAGGAGCTTAAGTACAGTCCCTGAGAGCAAAAGTTCCT 600
Db 214 MetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnAlaGlnLysValPro 233
QY 601 CATPACCAGAGAGCCCTGATCACTCATCTGTAGAAAACCAAAATCTCAAGAGATT 660
Db 234 HisAsnGlnArgLysProGlnSerThrHisLeuLeuGlnAsnGlnAsnThrGlnGluIle 253
QY 661 CCACCTGGCCATTAGCACTTCTCATCAGCAAGACAGACAGTGGCAAAACAGAGAT 720
Db 254 ProLeuAlaIleSerSerSerSerPheThrThrSerAsnHisSerValGlnAsnArgAsp 273
QY 721 CATAGAGAGAGCATGACGCCCTGTCTCTCAGAGATGCCAGCTCCAGCAGCTGCAGC 780
Db 274 HisArgGlyGlyMetGlnProCysSerProGlnArgCysGlnProAlaSerCysSer 293
QY 781 CCGCAGCAAAATATACGCCCCCTATAATACACAAGCTGGAGACACCTTCCCAAAAG 840
Db 294 ProHisGlnAsnIleLeuProLysIleThrSerThrArgProProPheProLysArg 313
QY 841 TCTGATGAAGAGTGTCCAGCACAAATGAATGATGATTCAGAGAAATACAGCCGACGCA 900
Db 314 SerAspArgLysAspValGlnHisAsnGlnIleIleGlyGlnLysArgGlnAla 333
QY 901 GTGAGAGAGCATTCATGAAGAGAGACAGAGATGATGTTCTTGGTCCGAGATTGTCC 960
Db 334 ValGlnGlnAlaPheMetLysGlnAsnLysAspLysSerPheLeuValArgAspCysSer 353
QY 961 ACAAAATCCAGAGAGAGCCCTATGTTGGCTGTGTTTATGAGACAAAGTCTACAT 1020
Db 354 ThrLysSerLysGlnGlnProLysValLeuAlaValPheLysGlnAsnLysValLysAsn 373
QY 1021 GTAAATAATCCGCTCTGAGAGAGAAATCAGAGATTGCCCTGGGAGAGAGACTCAGAGA 1080
Db 374 ValLysIleArgPheLeuGlnArgAsnGlnGlnPheAlaLeuGlnLysIleArgLys 393
QY 1081 GATGAGAGATTGATTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1128
Db 394 AspLysLysPheAspSerValGlnAspIleIleGlnHisIleLysAsn 409

```

RESULT 3

```

US-09-966-955A-6
; Sequence 6, Application US/09966955A
; Patent No. US20020155563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitley, Gena S.
; APPLICANT: Kanter, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-Length Human
; TITLE OF INVENTION: Signal-related Gene, MIST (Mast Cell Immunoreceptor
; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113US1

```

```

; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO: 6
; LENGTH: 353
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN MIST SPLICE VARIANT CLONE #12, TRANSLATED
; OTHER INFORMATION: AMINO ACID SEQUENCE
US-09-966-955A-6

```

Alignment Scores:

```

Pred. No.: 9,24e-136 Length: 353
Score: 1694.00 Matches: 317
Percent Similarity: 99.69% Conservative: 1
Best Local Similarity: 99.37% Mismatches: 1
Query Match: 83.86% Indels: 0
DB: Gaps: 0

```

US-09-856-061-3 (1-1129) x US-09-966-955A-6 (1-353)

```

QY 172 ATGAGAGAGCATGCGAGTGAATTAATTTACAGCCCGGCTATTAAGAACTGAA 231
Db 1 MetGlnGlnThrTrpGlnSerIleLysIleLeuProAlaArgProLysGlnSerGln 20
QY 232 TATGAGATACACACTATTTCAGAGTTGCAATGACACTCCCTCCGTTAGACACAGG 291
Db 21 TyrAlaAspThrHisIleTyLrPhelysValAlaMetAspThrProLeuProLeuAspThrArg 40
QY 292 ACCGTATTCCTCATTTGGACAGCCGACCTGAGACACAGACAGAGTGTGAAGAGTGCAC 351
Db 41 ThrSerIleSerIleGlyGlnProThrTrpAsnThrGlnThrArgLeuGlnArgValAsp 60
QY 352 AAACCATTTCCAGGAGCGTCAGAAAGCAAAACATTAAAGAGATGATCCGTAAGAAAG 411
Db 61 LysProIleSerLysAspValArgSerGlnAsnIleLysGlyAspAlaSerValArgLys 80
QY 412 AACAAATTCCTTATACACCTCCGCGGCTCTCATTAACCTTCGGAAGAGTCCAAACC 471
Db 81 AsnLysIleProLeuProProProArgProLeuIleThrLeuProLysLysIleGlnPro 100
QY 472 TTGCCCCCTGAGCGAGAGACAGAGCCACTTATCTCAGAGACACACTTCCAGAA 531
Db 101 LeuProProGlnProGlnSerSerArgProProLeuSerGlnArgHisIlePheProGln 120
QY 532 GTCCAGGGAATCCCACTCAGATTAAGCTTAAGGAGCTTAAGTGAAGTCTTGAAGCAGA 591
Db 121 ValGlnArgMetProSerGlnIleSerLeuArgAspLeuSerGlnValLeuGlnAlaGln 140
QY 592 AAATTCCTCATTAACAGAGAGAGACCTGATCACTCATCTGTAAGAAACCAAAATACT 651
Db 141 LysValProHisAsnGlnAlaGlyLysProGlnSerThrHisLeuLeuGlnAsnThr 160
QY 652 CAAGAGATTCACATTGCCATTAGCAGTCTTCATTCAGAGACCAAGCAACAGCTGTCAA 711
Db 161 GlnGlnIleProLeuAlaIleSerSerSerSerPheThrThrSerAsnHisSerValGln 180
QY 712 AACAGAGTCTATAGAGAGAGCATCAGCCCTGTTCTCTCAGAGATGCCAGCTCCAGCC 771
Db 181 AsnArgAspHisArgGlyGlyMetGlnProCysSerProGlnArgCysGlnProProAla 200
QY 772 AGCTGAGCCCTCAGAGAAATATATGCGCCATTAATTAACAGACACTGAGAGCAGCTTTC 831
Db 201 SerCysSerProHisGlnAsnIleLeuProLysIleGlnProLysIleGlnProPhe 220
QY 832 CCCAAAGAGTGTGATGAAGAGATGTCAGAGCAATGATGATGATGATGATGATGATGATGAT 891
Db 221 ProLysArgSerAspArgLysAspValGlnHisAsnGlnIleIleGlnLysIleGlnLysSer 240

```

US-09-864-761-38652
; Sequence 38652, Application US/09864761
; Patent No. US20020048763A1
; CURRENT INVENTION

? PRIOR APPLICATION NUMBER: US 60/180,312
 ? PRIOR FILING DATE: 2000-02-04
 ? PRIOR APPLICATION NUMBER: US 60/207,456
 ? PRIOR FILING DATE: 2000-05-26
 ? PRIOR APPLICATION NUMBER: US 09/632,366
 ? PRIOR FILING DATE: 2000-08-03
 ? PRIOR APPLICATION NUMBER: GB 24263.6

```

; TYPE: PRT
;; ORGANISM: Homo sapiens
; FEATURE:

```

US-09-856-061-3 (1-1129) x US-09-864-761-38652 (1-46)

Db 41 GlnPheAlaLeuGlyThr 46

; GENERAL INFORMATION:

CURRENT APPLICATION NUMBER: US/09/864,761

PRIOR APPLICATION NUMBER: PCT/US01/006622

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44023
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005599.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.47
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
; OTHER INFORMATION: SWISSPROT HIT: P25502, EVALUATE 3.70e+00
US-09-864-761-44023

```

```

Alignment Scores:
Pred. No.: 2,89e-10 Length: 40
Score: 214.00 Matches: 40
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.59% Indels: 0
DB: 10 Gaps: 0

```

US-09-856-061-3 (1-1129) x US-09-864-761-44023 (1-40)

```

QY 100 AGAAGCTTGTGCTGAGTCTGATGAGCAAGGCAAGCCAGTATGATGATGATGAC 159
      |||||||
Db 1 ArgstphealaaalvalleuaspclYAlalysclYHSSerapaspTyraasp 20

QY 160 CCTGAGCTTCGATGAGAGACATGCGAGTGAATTAATTTTACAGCCCGGCTATA 219
      |||||||
Db 21 ProgluenuargmetgluIutHtrpcolnserIleuysIleuProAlaargProIle 40

```

RESULT 6
US-09-864-761-44024

; Sequence 44024, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44024
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005599.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.47
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
US-09-864-761-44024

```

```

Alignment Scores:
Pred. No.: 2,89e-10 Length: 40
Score: 214.00 Matches: 40
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 10.59% Indels: 0
DB: 10 Gaps: 0

```

US-09-856-061-3 (1-1129) x US-09-864-761-44024 (1-40)

```

QY 98 AAGAACTTGTGCTGAGTCTGATGAGCAAGGCAAGCCAGTATGATGATGATG 157
      |||||||
Db 1 LysclurHleuargmetgluInsclYalHValmetHmet 20

QY 158 ACCCTGAGCTTCGATGAGAGACATGCGAGTGAATTAATTTTACAGCCCGGCTA 217
      |||||||
Db 21 ThrleuSerPheglYtrPlyalrghIsclYserHrleuysPheYrGlnProglYleu 40

```

RESULT 7
US-09-864-761-38653

; Sequence 38653, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:

; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04

```

; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38653
; LENGTH: 30
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005599.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4
US-09-864-761-38653

Alignment Scores:
Pred. No.: 3,22e-06 Length: 30
Score: 166.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 8,22% Indels: 0
DB: 10 Gaps: 0

US-09-856-061-3 (1-1129) x US-09-864-761-38653 (1-30)
QY 932 ATGTGAGTTTCTGTGTCGAGATGTTCCACAAATCCAGAGAGCCCATGTTTGG 991
Db 1 MetValValSerTPSerGluIleValIProGlnsProArgTlysSerProMetPheTrp 20
QY 992 CTGTGTTTATGAGAAAGCTATACATG 1021
Db 21 LeucysPheMetArgThrIlySerThrMet 30

RESULT 8
US-09-962-929-4
; Sequence 4, Application US/09962929
```

```

; Patent No. US20020115058A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.
; APPLICANT: Nielsen, Anne A.
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Diseases Associated wit
; TITLE OF INVENTION: Expression of Ptk3r1
; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/962,929
; PRIOR FILING DATE: 2000-09-22
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-962-929-4

Alignment Scores:
Pred. No.: 0.000147 Length: 724
Score: 150.50 Matches: 82
Percent Similarity: 37.04% Conservative: 48
Best Local Similarity: 23.36% Mismatches: 93
Query Match: 7,45% Indels: 128
DB: 10 Gaps: 17

US-09-856-061-3 (1-1129) x US-09-962-929-4 (1-724)
QY 406 AGAAGACAAAGATCTCTTACCACTCCGCGCTCTCATACACTTCGGAAGATAC 465
Db 80 LysIlySerProIleProThrProIlySerProArgPro-----ProArg----- 93
QY 466 CAACCTTCCCTCCCTGACCCGAGAGACAGAG----- 498
Db 94 --ProLeuProValAlaIleProGlySerSerIlyThrGluAlaAspValGluGlnAla 112
QY 499 -----CCACCTTATCTCAGAGACACACTTTCAGAAAGTC----- 534
Db 113 LeuThrIleuProAspLeuAlaGluGlnIleAlaProProAspIleAlaProProLeuLeu 132
QY 535 -----CAGGAATG----- 543
Db 133 IleIlyLeuValGluAlaIleGluIlyIlySerGlyLeuGlyCysSerThrIleuTyArgThr 152
QY 544 CCCAGTCGATTAAGCTTAAGGACTTAAGTCCCTGAGAGAGAAAGTCCAT 603
Db 153 GlnSerSerSerAsnLeuAlaGluIleuArgGlnIleuAspCysAsp---ThrProSer 171
QY 604 AACGAGAGAGCCCTGATCAACTCATCTGTAGAAACCA----- 645
Db 172 ValAspLeuGluMetIleAspValHisValLeuAlaAspAlaPheIlyArgTyIleuLeu 191
QY 646 -----ATACTCAAGAGATTCACCTGCGCATTAAGAGTTCT----- 681
Db 192 AspLeuProAsnProValIleProAlaAlaValIlySerGluMetIleSerLeuAlaPro 211
QY 681 ----- 681
Db 212 GluValGlnSerSerGluGluIlyTrIleGluLeuIlyLysLeuIleArgSerProSer 231
QY 682 -----TCA 684
Db 232 IleProHisGlnIlyTrIlePheIlyThrLeuGlnIlyTrIleuLeuIlyHisPhePheIlySer 251
QY 685 TTCGCGAAGCAACCAAGTGTGCAAAACAGATCATAGAGAGGATGAGCCCTGT 744
Db 252 GlnThrSerSerIlyAsnLeuIleuAsnAlaArg-ValIleuSerGluIlyIlePheSerProme 271
QY 745 TCTCTCAGA-----GATCGAGCCCTCCAGCCAGCTGCAAGCCCTCAGAA---ATA 795
Db 271 tleuPheArgPheSerAlaIleAspSerAspAsnThrGluAsnLeuIleIlyValIleGlu 291
```

QY 796 CTGCCCTATMAATACACAAGCTGAGAA-----CCACCTTTC 831
 |||||
 Db 291 uileuileiserthrclu--trpasnglurarginprolaleuproleprolys 310
 QY 832 CCCAAAAG-----TCTGATAGAAAAGATGTCGACGAC 864
 |||||
 Db 311 ProProlyserProthrThrValAlaAsnAsnglyMetAsnAsnAsnMetSerLeuGlnAsn 330
 QY 865 AATGATGTATACATGAGAAATACAGCCGACGAGCTGAGAGCATTCATGAGAGAG 924
 |||||
 Db 331 AlagIuTrpTyrTrpGlySerPheArgGlyValAsnGluLys---LeuArgAsp 349
 QY 925 AACCAAGATGATGATTCTTGTGTCGAGATGTTCCAAAATCCAAAGAACCCCTAT 984
 |||||
 Db 350 ThrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp---Tyr 368
 QY 985 GTTTGGCTGTGTTTATGAGAACAAAGTCACATGTAAA---ATCGCTTCTCGAG 1041
 |||||
 Db 369 ThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuLysIlePheHis 385
 QY 1042 AGAATACAGACGTTGCTCCCTGGGACAGACCTCAGAGAGATGAGAAGTTGATCAGTA 1101
 |||||
 Db 386 ArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheSerVal 401
 QY 1102 GAAGACATCATCGACACTACAGAAT 1128
 |||||
 Db 402 ValGluLeuIleAsnHisTyrArgAsn 410

RESULT 9

US-10-081-980B-1
 ; Sequence 1, Application US/10081980B
 ; Publication No. US20030041337A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pfizer Inc.
 ; APPLICANT: Gibbs, E. Michael
 ; APPLICANT: McNeish, John D.
 ; TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form Of
 ; FILE REFERENCE: Subunit Of PI-3 Kinase
 ; CURRENT APPLICATION NUMBER: US/10/081,980B
 ; CURRENT FILING DATE: 2002-08-23
 ; PRIOR APPLICATION NUMBER: 60/270,014
 ; PRIOR FILING DATE: 2001-02-20
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 688
 ; TYPE: PRT
 ; ORGANISM: mus musculus
 US-10-081-980B-1

Alignment Scores:

Pred. No.: 0.000195 Length: 688
 Score: 149.00 Matches: 76
 Percent Similarity: 32.76% Conservative: 96
 Best Local Similarity: 21.65% Mismatches: 140
 Query Match: 7.38% Indels: 17
 DB: 9 Gaps: 17

US-09-856-061-3 (1-1129) x US-10-081-980B-1 (1-688)

QY 421 CCTTACCACTCTCGGCTCTCATTAACACTTCCGAAGAATCAACCTTGCCTTC 480
 |||||
 Db 85 ProthrProlyserProArgPro-----ProArg-----ProLeuProVal 97
 QY 481 GAGCCGAGAGACAGCAGG-----CCACT 504
 |||||
 Db 98 AlarProGlySerSerLysThrGluAlaAspThrGluGlnGlnAlaLeuProLeuProAsp 117
 QY 505 TTATCTAGAGACACACTTTCCGAGAGTCAGAGGATGCCCATCATGATGATTAAGG 564
 |||||
 Db 118 LeuAlaGluGlnPheAlaProProAspVal---AlaProProLeuLeuLysLeuLeu 136

QY 565 GACTTAAGTAGAGTC-----CTTGAGCAGAAAAAGTTCTCATTAACCAAGAGAGCTT 618
 |||||
 Db 137 GluAlaIleGluLysLysGlyLeuGlnCysSerThrLeuTyrArgThrGlnSerSer 156
 QY 619 GAATCACTCATCTGTAA-----AACCAAAATCACTCAAGG 657
 |||||
 Db 157 AsnProAlaGluLeuArgGlnLeuLeuAspCysAspAlaAlaSerValAspLeuGluMet 176
 QY 658 ATTCCACTTGCCATTGATGAGCTTCTTCATTCACAGACAAAGCAACACAGCTGCAAAAACA 717
 |||||
 Db 177 IleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuAlaAspLeuProAsn 195
 QY 718 GATCATAGAGAGAGCATGACAGCCGTGTCCT----- 750
 |||||
 Db 196 -----ProValIleProValAlaValLysGlnLysGlnMetSer 208
 QY 751 -----CAGAGATGCCAGCTCCAGCCAGCTGC----- 777
 |||||
 Db 209 LeuAlaGluGlnLeuLeuInserProGluAspCysIleGlnLeuLeuLysLysLeuIleArg 228
 QY 778 -----AGCCCTCAGCAAAATATATCTGCCCTATAATAC----- 810
 |||||
 Db 229 LeuProAsnIleProHisGlnCysTrpLeuThrLeuGlnTyrLeuLeuLysHisPhePhe 248
 QY 810 ----- 810
 Db 249 LysLeuSerGlnAlaSerSerLysAsnLeuLeuAsnAlaArgValLeuSerGluIlePhe 268
 QY 810 ----- 810
 Db 269 SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrGluHisLeuIleLys 288
 QY 811 -----ACAAGCTGC----- 819
 |||||
 Db 289 AlaIleGluIleLeuLysSerThrGluTyrPasnGluArgGlnProAlaProAlaLeuPro 308
 QY 820 ---AGACCCTTTCCCTC-----AAAAGTCTGATGAAGAAGATGC 858
 |||||
 Db 309 ProLysProProLysProThrThrValAlaAsnAsnSerMetAsnAsnMetSerLeu 328
 QY 859 CAGCACATATATGTRACATGAGATACAGCCGACGAGCTGAGAGGCAATCATG 918
 |||||
 Db 329 GlnAspAlaGluTyrPyrTrpGlyAspIleSerThrGluGluValAsnGluLys---Leu 347
 QY 919 AAGGAGACAAGAGATGTTCTTGTGTCGAGATGTTCCAAAATCCAAAGAGAAG 978
 |||||
 Db 348 ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp 367
 QY 979 CCTTATGTTTGGCTGTGTTTATGAGAACAAAGCTACATGTAAA---ATCCGCTTC 1035
 |||||
 Db 368 ---TyrThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIle 383
 QY 1036 CTGAGAGAGATACAGAGCTTGCCTGGGACAGAGACTCAGAGAGATGGAAGTTGAT 1095
 |||||
 Db 384 PheHisArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheAsn 399
 QY 1096 TCAGTAGAAGCATCATCGAAGCACTACAGAAT 1128
 |||||
 Db 400 SerValValGluLeuIleAsnHisTyrArgAsn 410

RESULT 10

US-10-081-980B-3
 ; Sequence 3, Application US/10081980B
 ; Publication No. US20030041337A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Pfizer Inc.
 ; APPLICANT: Gibbs, E. Michael
 ; APPLICANT: McNeish, John D.
 ; TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form
 ; FILE REFERENCE: Subunit Of PI-3 Kinase
 ; CURRENT APPLICATION NUMBER: US/10/081,980B

```

:
: CURRENT FILING DATE: 2002-08-23
: PRIOR APPLICATION NUMBER: 60/2270, 01/4
:
: PRIOR FILING DATE: 2001-02-20
:
: NUMBER OF SEQ ID NOS: 12
:
: SOFTWARE: PatentIn version 3.1
:
: SEQ ID NO 3
:
:   LENGTH: 724
:   TYPE: PRY
:
: ORGANISM: mus musculus
:
: US-10-081-980B-3

```

Pred. No.:	0.000197	Length:	72
Score:	149.00	Matches:	76
Percent Similarity:	32.76%	Conservative:	39
Best Local Similarity:	21.65%	Mismatches:	96
Query Match:	7.38%	Indels:	140
DB:	9	Gaps:	17

US-09-856-061-3 (1-1129) x US-10-081-980B-3 (1-724)

QY	421	CCCTTACCACTCTCTCGGCTCTCATTAACACTTCGGAAGAAGTACCAACCTTGCCCTC	480
Db	85	ProthrProLysProArgPro-----ProArg-----ProLeuProVal	97
QY	481	GAGCCGAGACAGCAGCAG-----	504
Db	98	AlaProLysSerLysThrGluAlaAspThrGluGlnGlnAlaLeuProLeuProAsp	117
QY	505	TTATCTCAGACACACACTTTCGGAAGTCCAGGAATGGCCGATCGATACCTTAAG	564
Db	118	LeuAlaGluGlnPheAlaProProAspVal---AlaProProLeuLeuLysLeuLeu	136
QY	565	GACTTAAGTGAGTTC-----CTTGAAGCAGAAAAAGTTCCTATAACCGAAGAACCT	618
Db	137	GluValAlaIleGluLysGlyLeuGluLucySerThrLeuYrArgThrGlnSerSer	156
QY	619	GAATCACTCATCTGTTAGAA-----AACCAAAATACTCAAGAG	657
Db	157	AsnProAlaGluLeuArgGlnLeuLeuAspCysAspAlaAlaSerValaAspLeuGluMet	176
QY	658	ATTCCACTTGCATATAGCAGTCTTCATTCACGACAGCAACACACAGTGTGCAAAACAG	717
Db	177	IleAspValHisValLeuAlaAspAlaPheLysArgGlyLeuAlaAspLeuProAsn---	195
QY	718	GATCATGTAGAGGAGCATGACCCCTGTTCTCT-----	750
Db	196	-----ProValIleProValAlaValTyrAsnGluMetMetSer	208
QY	751	-----CAGAGNTGCCAGCCTCCAGCCAGCAGCTGC-----	777
Db	209	LeuAlaGlnIleuLeuGlnInsProGluAspCysIleGlnLeuLeuLysLysLeuIleArg	228
QY	778	-----AGCCCTCAGCAAAATATACGTCCCTATATAAATAC-----	810
Db	229	LeuProAsnIleProHisGlnCysThrPheThrLeuGlnIleuLeuLysHisPhePhe	248
QY	810	-----	810
Db	249	LysLeuSerGlnAlaSerSerLysAsnLeuLeuAsnAlaArgValLeuSerGluIlePhe	268
QY	810	-----	810
Db	269	SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrGluHisLeuIleLys	288
QY	811	-----ACAACCTG-----	819
Db	289	AlaIleGluIleLeuLysSerThrGluTyrAsnGluArgGlnProAlaProAlaLeuPro	308
QY	820	-----AGACACACTTTCCTC-----	858
Db	309	ProLysProProLysProThrThrValAlaAsnAsnSerMetCysAsnMetSerLys	328

QY	859	CAGCAGCATGATGATGTCATCATGAGGATATACAGCCAGCAGCATGTGGAAAGGCATTCATG	918
Db	329	GlnAspAlaGluTyrPtyrTyrPtyrPglYAspIleSerArgGluGluValAlaSnGluYs---	347
QY	919	AAGGAGAAACAAGATGGTAGTTCTTGTCGCGAGATGTTCCACAAATCCAGAGAGAG	978
Db	348	ArgAspThrAlaAspGlyThrPheLeuValAlaArgAlaSerThrIlysmethIscIyAsp	367
QY	979	CCCTATGTTTGGCTGGTGTTTATGAGACAAAGTCTACAACTGTAA---ATCCGCTTC	1033
		::	
Db	368	--TyrThrLeuThrIleu-----ArgIlyGlyGlyIysAlaSnIlyLeuIleuIle	383
QY	1036	CTCGAGAGGAATCGACATTTGCCCTGGGAGACAGACTCAGAGAGATGAGAGATTGAT	1095
		:: :: :: :: :: :: :: :: :: ::	
Db	384	PheHisAlaGAspGlyIysTyrGlyPheSerAspProIeu-----ThrPheAsn	399
QY	1096	TCAGTAAAGACATCATCGAACAATCAACAT	1128
		:: ::	
Db	400	SerValValGluLeuIleAlaSnHisTyrArgAsn	410

RESULT 11
US-09-962

US-09-962-929-2
; Sequence 2, Application US/09962929
; Patent No. US20020115058A1
GENERAL INNOVATIONS

; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.

; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Diseases Associated with
; TITLE OF INVENTION: Expression of Plk3r1

```

; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,929
; CURRENT FILING DATE: 2001-08-24

```

; COMMENT FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: US 09/668,644
 ; PRIOR FILING DATE: 2000-09-22

```

; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: Patentin version 3.1

```

; SEQ ID NO 2
; LENGTH: 724

```

; TYPE: PRT
; ORGANISM: Mus musculus
; NC-00-063-030

```

Alignment Scores:	
Pred. No.:	0.000292
Score:	147.00
Percent Similarity:	32.48%
Best local Similarity:	21.65%
Query Match:	7.28%
DB:	10
Length:	724
Matches:	76
Conservative:	38
Mismatches:	97
Indels:	140
Gaps:	17

US-09-856-061-3 (1-1129) x US-09-962-929-2 (1-724)

OY	421	CCCTTACCAACTCGTCGGGCGCTTCATTAACACTTCCGAAGAAAGTACCAACCTTGCCGCCCT	480
Db	85	ProthrProLysProlAlaGpro-----ProlArg-----ProlProVal	97
OY	481	GAGCCGAGAGACAGCAGC-----GCACCT	504
Db	98	AlaProGlySerSerylStHgluaIalaSpThrcluglnGlualenProleuproasp	117
OY	505	TTATCTCAGAGCACACCCTTTCCAGAAGTCCAGGGAATGCCCACTCATGATTAACTTAAG	564
Db	118	LeuAlaIuglnPhealaproProAspVal--AlaProProleuLeuileLysleuLeu	136
OY	565	GACTTAAGTAGGRC-----CTTGAAGCAGAAAATTCCTCATTAACAGCAGAGAAGCCT	618
Db	137	GlnAlaIlegluLysLysgLyLeuGlucySerThrLeuYrnrghTrghInserSerSer	156
OY	619	GAATCAACTATCTGTGTAAG-----AACCAAAATCACTAAG	657
Db	157	AsnProIacIugluendArvglnLeuAspCyasApAlaIaserValaspLeuGlnmet	176

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:15:56 ; Search time 35.2565 Seconds
(without alignments)
6156.924 Million cell updates/sec

Title: US-09-856-061-3
Perfect score: 2020
Sequence: 1 ttccagacttcagctgcgc.....catcgaactacaagaatt 1129

Scoring table:
BL0SUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283224 segs, 96134422 residues
Total number of hits satisfying chosen parameters: 566448

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODL=frame+_{n2p}.model -DEV=xlp
-O=/sgn2.1/USPTO_spool/US09856061/runat_21042003.113019.3758/app_query.fasta_1.3150
-DB=PIR.73 -QFMT=fastan -SUFEX=n2p.rpr -MINMATCH=0.1 -DOOPCL=0 -DOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=bl0sum62 -TRANS=human4.0.cdi -LIST=45
-DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=fto -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061.ecgn.1.1.151.e/runat_21042003.113019.3758 -NCPU=6 -ICPU=3
-NO_XLPYX -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV=TIMEOUT-120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXt=7

Database :
1: PIR.73:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	332.5	16.5	533	2	A56110
2	332.5	15.6	533	2	tyrosine phosphor
3	154.5	7.6	724	2	A38749
4	150.5	7.5	724	2	3-phosphatidylinos
5	149	7.4	724	2	phosphatidylinos
6	143	7.1	659	2	SH3 binding protei
7	140	6.9	559	2	protein-tyrosine k
8	140	6.9	659	1	protein-tyrosine k
9	139	6.9	663	1	gag-abl polypeptid
10	138.5	6.8	841	1	protein-tyrosine-p
11	136.5	6.5	485	2	hypothetical colle
12	131.5	6.5	723	2	3-phosphatidylinos
13	131	6.3	728	2	phosphoinositide-3
14	126.5	6.3	728	2	H59435

15	125	6.2	377	2	S08636	nck protein - huma
16	125	6.2	585	2	A46209	protein-tyrosine-p
17	125	6.2	593	1	JM0805	protein-tyrosine-p
18	125	6.2	595	1	A55651	protein-tyrosine-p
19	125	6.2	597	1	A53593	protein-tyrosine-p
20	124.5	6.2	2649	2	T51023	hypothetical prote
21	123.5	6.1	1196	2	T23832	protein-tyrosine k
22	123	6.0	461	2	A57463	p55PIK - mouse
23	122	6.0	822	1	TYHUPF	transcription init
24	121.5	6.0	1893	1	A40262	hypothetical prote
25	121	6.0	1171	1	T12956	protein M01A8.2 (1
26	120.5	6.0	893	2	G88551	protein-tyrosine-p
27	120	5.9	593	2	JC5167	probable serine/th
28	120	5.9	939	2	S28394	nonphosphotyrosine
29	120	5.9	996	2	T47518	Shb-like adapter p
30	119	5.9	480	2	JC5552	protein-tyrosine k
31	119	5.9	620	1	S33253	probable transcript
32	118.5	5.9	1392	2	T51947	hypothetical prote
33	118	5.8	546	2	S52314	hypothetical prote
34	117.5	5.8	289	2	T52354	hypothetical prote
35	117.5	5.8	667	2	T27672	hypothetical prote
36	117.5	5.8	1357	2	T16860	hypothetical prote
37	116.5	5.8	1162	2	T49191	hypothetical prote
38	116	5.7	545	2	S52313	protein-tyrosine k
39	116	5.7	1172	2	T00065	hypothetical prote
40	115.5	5.7	3534	2	T42567	legument protein 2
41	115	5.7	675	2	S60612	protein-tyrosine k
42	115	5.7	796	2	E96654	hypothetical prote
43	115	5.7	1290	2	A36466	1-phosphatidylinos
44	114.5	5.7	625	1	A43030	protein-tyrosine k
45	114.5	5.7	630	1	T01380	protein-tyrosine k

ALIGNMENTS

RESULT 1
A56110
tyrosine phosphoprotein SLP-76 - human
C:Species: Homo sapiens (man)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
C:Accession: A56110
R:Jackman, J.K.; Mochly, D.G.; Sun, Q.; Tanemoto, M.; Turk, C.W.; Peltz, G.A.; Koretz
J. Biol. Chem. 270, 7029-7032, 1995
A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated wit
A:Reference number: A56110; MUID:95221345; PMID:7706237
A:Accession: A56110
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <YAC>
A:Cross-references: GB:U20158; NID:9806765; PIDN:AAC50135.1; PID:9806766
C:Genetics:
A:Gene: GDB:ICP2; SLP-76
A:Cross-references: GDB:1230199; OMIM:601603
A:Map position: 5q33.1-5pter
C:Superfamily: SH2 homology
C:Keywords: Phosphoprotein
F:422-520/Domain: SH2 homology <SH2>

Alignment Scores:
Pred. No.: 1,32e-18
Score: 332.50
Percent Similarity: 36.24%
Best Local Similarity: 25.28%
Query Match: 16.46%
DB: 2
Gaps: 16

US-09-856-061-3 (1-1129) x A56110 (1-533)
QY 58 GGCAGACAGCAGATGATGAGCAAGACCTCTGACTGGAAGAACTTGTCGAGTC 117
DB 126 GATGATGAGCAAGACAGCAGTGTGATGATGATGATGATGATGATGATGATGAT 139
QY 118 CTGATGAGCAAGACAGCAGTGTGATGATGATGATGATGATGATGATGATGATGAT 177

```

Db 140 GLUAsp-----AspAlaAspTyrGluProProSerAsnAsp153
QY 178 GAGACATGGCAGTGGATTAAATTTTACACCGCGCTATAAAGATCGATATGCA 237
Db 154 GluAlaLeuGlnAsn---SerIleuProAlaLysProPheProAsnSer----- 169
QY 238 GATACACACTATTTCAGAGTTTCAATGAGACACTCCCTTCGTTAGACACGAGACTCT 297
Db 170 AsnSerMetTyr----- 173
QY 298 ATTCACATTTGGACACCGACCTGGACACACAGAGTTTGAAGAAGTGGACAAACC 357
Db 174 ----- 177
QY 358 ATTTCAGAGGACGTGACAGCCCAAAACATTAAAGAGATGCATCCGTAAAGAAACAG 417
Db 178 ProSerGluLysThrProGlnGlnPro----- 186
QY 418 ATTCCTTTACACCTCTCGGCTCTCATACACTCCG-----AAGAG 462
Db 187 ---ProValProProGlnArgProMetAlaAlaLeuProProProProAlaGluArgAsn 205
QY 463 TACCAACCTTGGCCCT-----GAGCCGAGACAGCAGG----- 498
Db 206 HisSerProLeuProProProGlnThrAsnHisGluLysProSerArgSerArgAsnHis 225
QY 498 ----- 498
Db 226 LysThrAlaLysLeuProAlaProSerIleAspArgSerThrLysProProLeuAspArg 245
QY 499 ----- 513
Db 246 SerIleuAlaProPheAspArgGluProPheThrLeuGluLysLysProProPheSerAsp 265
QY 514 AGACACACCTTTTCA-----GAAGTCAG 537
Db 266 LysProSerIleProAlaGluArgSerLeuGluHisLeuProLysIleGlnLysPro 285
QY 538 GGAATGCCACATGATAGACTTAAGGACTTAAGAGTCTTGAAGAGAAAGTT 597
Db 286 ProLeuProProThrThrGlnArgHisGluArgSerSerProLeuProGluLysLysPro 305
QY 598 CCF-----CAT 603
Db 306 ProValProLysHisGluTyrGluProAspArgGluAsnAspGluAspValHis 325
QY 604 AACACAGAGAAAGCTGAATCAATCATCTGTAGAAACCAAAATACT----- 651
Db 326 GlnArgProLeuProGlnProAlaLeuLeuProMetSerSerAsnThrPheProSerArg 345
QY 652 -----CAAGAGTTCCACTTGCCTTATGACACTTTCATTCACG 690
Db 346 SerThrLysProSerProMetAsnProLeuProSerSerHisMetProGluAlaPheSer 365
QY 691 ACAAGCAACACAGTGTGCAAAACAGATCATAGAGAGACATGACCCCTGTCTCT 750
Db 366 GluSerAsnSerPhe-----ProGlnSerAlaSerLeuProTyrPheSer 382
QY 751 CAG-----AGATGCCAGCTCCAGCCAGCTGACGCCCTACGAAATATATACGCCCTAT 804
Db 383 GlnGluProSerAsnArgProProIleArgAlaGluGluLysArgAsnPheProLeu 402
QY 805 AAATACACAAGCTGGAGACACTTTCCCAAAAGCTGTGATAGAAAGATGTCAGAC 864
Db 403 Pro---AsnLysProArgProProSerPro-----AlaGluGluGluAsnSerLeuAsn 419
QY 865 AATGATGTGATGAGAGAAATACAGCGCCAGCAGTGAAGACATTCATGAAGAG 924
Db 420 GluGlnThrTyrValSerTyrIleThrArgProGluAlaLysAlaLeuArgLysIle 439
QY 925 AACAGAGATGATGTTCTGTGTCGAGATGTTCCCAAAATCCAAAGAGACCTAT 964

```

```

Db 440 AsnGlnAspGluThrPheLeuValArgAspSerSerLysLysThrThrAsnProTyr 459
QY 985 GTTTGGCTGTCTTTTATGAACAAGCTTACATGTAATAATCCGTTCTGAGAGG 1044
Db 460 ValLeuMetValLeuTyrLysAspLysValTyrAsnIleGlnIleArgTyrGlnLysGlu 479
QY 1045 AATCAGCAGTTTGGCCGGGAGCAGCTGAGAGAGATGAGAGTGTGATTCAGTGA 1104
Db 480 SerGlnValTyrLeuLeuGluTyrGlyLeuArgGluLysGluAspPheLeuSerValSer 499
QY 1105 GACATCATCGAACACTACAG 1125
Db 500 AspIleIleAspTyrPheArg 506

```

RESULT 2

B56110
 Lysine phosphoprotein SLP-76 - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
 C:Accession: B56110
 R:Jackson, J.K.; Monto, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretz
 J. Biol. Chem. 270, 7029-7032, 1995.
 A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with
 A:Reference number: A56110; MUID:95221345; PMID:7706237
 A:Accession: B56110
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-533 <JNC>
 A:Cross-references: GB:U20159; NID:9806767; PIDN:AAC52189.1; PID:9806768.
 C:Superfamily: SH2 homology
 C:Keywords: phosphoprotein
 F:422-520/Domain: SH2 homology <SH2>

Alignment Scores:

Pred. No.:	3.03e-17	Length:	533
Score:	315.50	Matches:	121
Percent Similarity:	40.94%	Conservative:	44
Best Local Similarity:	30.02%	Mismatches:	169
Query Match:	15.62%	Indels:	69
DB:	2	Gaps:	18

US-09-856-061-3 (1-1129) x B56110 (1-533)

```

QY 58 GCCGATGCCAGAGATGACAAAGCCTCTTCTAGACTGGAAAGAACTTGTGTCAGTC 117
Db 126 GlyAspTyrGluSerProAsn-----GluGluGluGlnAlaLeuVal 139
QY 118 CTGATGAGACAAAGGCCACAGTATGATGACTATGATGACCCCTGAGCTTCGAGTGA 177
Db 140 AspAspAlaIle-----AspTyrGluProProProSerAsnGln 153
QY 178 GAGACATGGCAGTGGATTAAATTTTACACCGCGCTATA-----AAGAAATCGAA 231
Db 154 GluAlaLeuGlnSer---SerIleLeuProProAsnSerPheHisAsnThrAsnSerMet 172
QY 232 TATGCAAGT---ACACACTATTTCAGTGTGATGACACTCCCTT---CCGTTAGAC 285
Db 173 TyrIleAspArgProProThrGluLysValSerGlnGlnProValProProLeuArg 192
QY 286 ACAGAGACTCTATTCATTCATGATGACAGCCAGCTGGAACACACAGAGAGTTGGAAGA 345
Db 193 ProLysProAlaLeuPro---ProLeuProThrGluArgAsnHisSerProLeuSerPro 211
QY 346 GTGAGCAAAACCATTTTCAGAGGAGCTGACAGCCCAAAACATTAAAGAGATGCATCCGTA 405
Db 212 ProHisProAsnHisGluGluProSerArgSerGlyAsnAsnLys----- 226
QY 406 AGAAAGACAAGATTCCTTTACCA-----CCTCCTCGGCTCTC 444
Db 227 ---ThrAlaLysLeuProAlaProSerIleAspArgSerThrLysProProLeuAspArg 245
QY 445 ATTAACACTTCCGAAGAAGTACCAACCTTG-----CCGCTGAGCCGGAG 489

```

Db 246 SerLeuAlaProLeuAspArgGluProPheIleLeuGlyLysLysProProPheSerAsp 265
 QY 490 AGCAGAGCGCCACCTTATCTCAGAGACACACTTTCAGAAAGTCCAG-----GGAAATG 543
 Db 266 LysProSerAlaProLeuGlyArgGluHis---LeuProLysIleGlnLysProProLeu 284
 QY 544 CCCAGTCAGATTAAGCTTAAGGAGCTTAAGTGGCTCTT-----GAAGCA 588
 Db 285 ProProAlaMetAspArgHisGluArgAsnGluArgLeuLysProValThrThrArgLys 304
 QY 589 GAAAAAGTTCCTCATTAACCGAGAGAGCCCTGAATCAACTCATCTGTTAGAA----- 639
 Db 305 ProSerValProArgHisGlyArgGlyProAspArgArgLysAsnAspGluAspAspVal 324
 QY 640 AACCAAAATATCTCAAGAGATTCCTGACCTGACAGTTCCTCATTCAGACAGCAAGAAC 699
 Db 325 HisGlnArgProLeuProGlnProSerLeuProSerMetSerSerAsnThrPheProSer 344
 QY 700 CACAGTGTGCA----- 711
 Db 345 ArgSerValGlnProSerSerLysAsnThrPheProLeuAlaHisMetProGlyAlaPhe 364
 QY 712 AACAGATCATATAGAGAGGACGACCCCTGTTCT-----CCTCAGAGATGCCAGCT 765
 Db 365 SerGluSerAsnIleGlyPheGlnGlnSerAlaSerLeuProProTyrrPheSerGlnGly 384
 QY 766 CCGAGCCAGTCGACCCCT-----CACGAAATATACTGCCCTTAATAATACACAGC 816
 Db 385 ProGlyAsnArgProProLeuArgSerGluGlyArgAsnLeuProLeuProValProAsn 404
 QY 817 TGGAGACACCTTCCCAAAAGGTCTGATAGAAAGATGTCAGACCAATGAATGTATC 876
 Db 405 ---ArgProGlnProProSerProGlyGluGlnGlnThrProLeuAspGluIleTyrr 423
 QY 877 ATTGAGCAATACAGCCGCGCAGAGCAGTGGAGAGAGCATTCATGAAGGAGAACAGAGTGT 936
 Db 424 ValSerTyrrIleThrArgProGluAlaGluAlaLeuArgLysIleAsnGlnAspGly 443
 QY 937 AGTTTCTGTCGCGAGATGTTCCACAAATTCAGAGAGACCCCTATGTTGGCTGTG 996
 Db 444 ThrPheLeuValArgAspSerSerLysLysThrAlaAsnAsnProTyrrValLeuMetVal 463
 QY 997 TTTATGAGAACAAAGTCTCAATGTAAATAATCGCTTCTCGAGAGAAATCAGACACTT 1056
 Db 464 LeuTyrlLysAspLysValTyrrAsnIleGlnIleArgTyrrGlnGlnLysSerGlnValTyrr 483
 QY 1057 GCCCTGGGCGACAGACTCAAGAGATGAGAGATTGATTTCAGTGAAGACATCATCGAA 1116
 Db 484 LeuLeuGlyThrGlyLeuArgGlyLysGluAspPheLeuSerValSerAspIleIleAsp 503
 QY 1117 CACTACAG 1125
 Db 504 TyrPheArg 506
 RESULT 3
 A38749
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85K chain alpha - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Feb-1992 #sequence, revision 14-Feb-1992 #text, change 05-Nov-1999
 C:Accession: A38749
 R:Otsu, M.; Hiles, I.; Goult, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, A.
 Cell 65, 91-104, 1991
 A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine ki
 A:Reference number: A38749; MIMD:91191567; PMID:1707345
 A:Accession: A38749
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-724 <OTS>
 A:Cross-references: GB:M61746; GB:M61745; NID:J163476; PIDN:AAA79511.1; PID:J163477
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:333-428/Domain: SH2 homology <SH2A>

F:624-718/Domain: SH2 homology <SH2>
 Alignment Scores:
 Pred. No.: 0.000237 Length: 724
 Score: 154.50 Matches: 81
 Percent Similarity: 37.21% Conservative: 47
 Best Local Similarity: 23.55% Mismatches: 102
 Query Match: 7.65% Indels: 114
 DB: 2 Gaps: 14
 US-09-856-061-3 (1-1129) x A38749 (1-724)
 QY 406 AGAAGAACAAAGATTCCTTACCACTCTCGGCTCTCATTAACACTTCGAAAGACTAC 465
 Db 80 LysLysIleSerProProThrProLysProArgProProArgProLeuProValAlaPro 99
 QY 466 CAACCTTGGCCCTTGAGCCGAGAGACAGAG-----CCACTTATCT 510
 Db 100 GlyProSerLysThrGlnAlaAspSerGluGlnGlnAlaSerThrLeuProAspLeuAla 119
 QY 511 CAGAGACACACCTTCCAGAAATC----- 534
 Db 120 GluGlnPheAlaProProAspValAlaProProLeuLeuIleLysLeuValGluAlaIle 139
 QY 535 -----CAGGGAATGCCAGTCAGATTAAGCTTAAG----- 564
 Db 140 GluLysLysGlyLeuGluLysSerThrLeuTyrrArgThrGlnSerSerSerAsnProAla 159
 QY 565 GACTTAAGTAGAGTCTTGAAGCAGAAAGACTCT----- 600
 Db 160 GluLeuArgGlnLeuLeuAspPcyAspThrAlaSerLeuAspLeuGluMetPheAspVal 179
 QY 601 CATTAACAGAGAGACCTGATCAATCAATCATTGTTAGAAACCAAAATATCAAGAGAT 660
 Db 180 HisValLeuAlaAspAlaPheLysArgTyrrLeuLeuAspLeuProAsn---ProValIle 198
 QY 661 CCACCTGCCATTACAGACTCT----- 681
 Db 199 ProValAlaValSerSerGluLeuIleSerLeuAlaProGluValGlnSerSerGlu 218
 QY 681 ----- 681
 Db 219 TyrrIleGlnLeuLeuLysLeuIleArgSerProSerIleProHisGlnTyrrPhe 238
 QY 682 -----TCATTACAGACAGACACACAGT 705
 Db 239 ThrLeuGlnTyrrLeuLeuLysHisPhePheLysLeuSerGlnThrSerLysAsnLeu 258
 QY 706 GTGCMAAACAGAGATCATAGAGAGAGCATGCGCCCTGCTCCAGAGATGCCAGCT 765
 Db 259 LeuAsnAlaArgValLeuSerGluLeuPheSerProLeuLeuPheArgPheProAla 278
 QY 766 CAGCCAGCTGACGCCCTCACGAAATATACTGCCCTTAATAATC-----ACAAGC 816
 Db 278 laser-SerGluAsnThrGlnHisLeuIleLysIleIleGlnIleLeuIleSerThrGlu 297
 QY 817 TGGAGA-----CCACCTTCCCAAAAG----- 840
 Db 298 ThrAsnGluArgGlnProAlaProAlaLeuProProLysProProLysProThrVal 317
 QY 841 -----TCGTATGAAAGAGATGTCAGACCAATGAATGGTATCGAGAA 885
 Db 318 AlaAsnAsnGlyMetAsnAsnAsnMetSerLeuGlnAspAlaGluIleTyrrTyrrGlyAsp 337
 QY 886 TACAGCCGCGCAGAGTGAAGAGCATTCATGAAGGAGAACAGATGATGTTCTG 945
 Db 338 IleSerArgGluGluValAsnGluLys---LeuArgAspThrAlaAspGlyThrPheLeu 356
 QY 946 GTCCAGATGTTCCACAAATTCAGAGAGACCCCTGTTGGCTGTTTATGAG 1005
 Db 357 ValArgAspAlaSerThrLysMetHisGlyLys---TyrrThrLeuThrLeu----- 372
 QY 1006 AACCAAGTCTCAATGTAAATA---ATCCGTTCTGAGAGAAATACAGACTTGGCCCTG 1062

```

Db 373 ArglySGlyGlyAsnAsnLysLeuIlePheHisArgAspLysTyrGlyPhe 392
QY 1063 GGGACAGGAGCTCAGAGAGATGAGAGTTGATTCAGTACATCCAGACCTAC 1122
Db 393 SerAspProLeu-----ThrPheAsnSerValGluLeuIleAsnHisTyr 408
QY 1123 AAGAAAT 1128
Db 409 ArgAsn 410

RESULT 4
A:38748
3-phosphatidylinositol kinase (EC 2.7.1.-) 85k chain - human
C:Species: Homo sapiens (man)
C:Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 16-Jul-1999
C:Accession: A38748; S28402
R:Skolnik, E.Y.; Margolis, B.; Mohammadi, M.; Lowenstein, E.; Fischer, R.; Drepps, A.; U
Calli 65, 83-90, 1991
A:Title: Cloning of p13 kinase-associated p85 utilizing a novel method for expression/cl
A:Reference number: A38748; MUID:91191565; PMID:1849461
A:Accession: A38748
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-724 <SKD>
A:Cross-references: GB:M61906
R:Panayotou, G.; Bax, B.; Gout, I.; Federlisch, M.; Wroblewski, B.; Dhand, R.; Fry, M.J.
EMBO J. 11, 4261-4272, 1992
A:Title: Intersection of the p85 subunit of PI 3-kinase and its N-terminal SH2 domain wit
A:Reference number: S28402; MUID:93049176; PMID:1330535
A:Accession: S28402
A:Status: Preliminary
A:Molecule type: Protein
A:Residues: 301-311,315-319;424-439 <PAN>
A:Gene: GDB:PIK3R1
A:Cross-references: GDB:127604; OMIM:171833
A:Map position: 5q12-5q13
C:Superfamily: SH2 homology
C:Keywords: phosphotransferase
F:333-428/Domain: SH2 homology <SH2A>
F:624-718/Domain: SH2 homology <SH2>

Alignment Scores:
Pred. No.: 0.000496 Length: 724
Score: 150.50 Matches: 82
Percent Similarity: 37.04% Conservative: 48
Best Local Similarity: 23.36% Mismatches: 93
Query Match: 7.45% Indels: 128
Db: 2 Gaps: 17

US-09-856-061-3 (1-1129) x A38748 (1-724)
QY 406 AGAAGACAGAGATTCCTTACACCTCGGCTCATACACTTCGAGAGAGTAC 465
Db 80 LysLysIleSerProThrProLysProArgPro-----ProArg----- 93
QY 466 CAACCTTGGCCCTGACCGGAGACGACGAG----- 498
Db 94 ---ProLeuProValAlaLapProGlySerSerLysThrGluAlaAspValGluGlnAla 112
QY 499 -----CCACCTTATCTTCAGAGACACACCTTCCAGAAAGTC----- 534
Db 113 LeuThrLeuProAspLeuAlaGluGlnPheAlaProAspIleAlaProProLeuLeu 132
QY 535 -----CAGGGAATG----- 543
Db 133 IleLysLeuValGluAlaIleGluLysLysGlyLeuGluCysSerThrLeuTyrArgThr 152
QY 544 CCCAGCTAGATTAAGGAGACTTAAGTGAAGTCTTGAAGCAGAAAAGTTCCTCAT 603
Db 153 GlnSerSerSerAsnLeuAlaGluLeuArgGlnLeuLeuAspCysAsp---ThrProSer 171

```

```

QY 604 AACCAAGAGAGCGCTGATCATCTATCTTTAGAAACCA----- 645
Db 172 ValAspLeuGluMetIleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuLeu 191
QY 646 -----AATACTCAAGAGATTCCTGCTTCCATTAGCAGTCT----- 681
Db 192 AspLeuProAsnProValIleProAlaIleValTyrSerGluMetIleSerLeuAlaPro 211
QY 681 ----- 681
Db 212 GluValGlnSerSerGluGluTyrIleGlnLeuLeuLysLeuLeuIleArgSerProSer 231
QY 682 -----TCA 684
Db 232 IleProHisGlnTyrTrpLeuThrIleGlnTyrLeuLeuLysHisPhePheLysLeuSer 251
QY 685 TTCACGACAGCAACCAACAGTGTGCAAAACAGATCATAGAGAGCCATGACCCCTT 744
Db 252 GlnThrSerSerLysAsnLeuLeuAlaArg-ValLeuSerGluIlePheSerProkoe 271
QY 745 TCTCCCTCAGA-----GATGCCAGCCTCCAGCCAGCTGACGCCCTCAGCAAA---ATATA 795
Db 271 LeuPheArgPheSerAlaAlaSerSerAspAsnThrGluAsnLeuIleLysValIleG1 291
QY 796 CTGCCCTATTAATACACAAAGCTGAGA-----CCACCTTTC 831
Db 291 uIleLeuIleSerThrGlu--TrpAsnGluArgGlnProAlaProAlaLeuProProLys 310
QY 832 CCCAAAGG-----TCTGATAGAAAGGATGTCACAGAC 864
Db 311 ProProLysProThrThrValAlaAsnAsnGlyMetAsnAsnMetSerLeuGlnAsn 330
QY 865 AATGATGTGATCATGAGATATACAGCCGCCAGGACGAGAGGAGCATTCGAAGAG 924
Db 331 AlaGluTrpTyrTrpGlyAspIleSerArgGluGluValAsnGluLys---LeuArgAsp 349
QY 925 AACCAAGATGAGTATTTCTGTGTCGAGATTTGCCAATAATCCAGAGAGCCCTAT 984
Db 350 ThrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp---Tyr 368
QY 985 GTTTGGCGTGTGTTTATGAGAACAAAGTTCATGTGAAA---ATCCGCTTCCTGGAG 1041
Db 369 ThrLeuThrLeu-----ArgLysGlyLysAsnLysLeuIleLysIlePheHis 385
QY 1042 AGCAATCAGACAGTTCGCCCTGGGACAGACTCAGAGAGATGAGAGATTCATTCAGA 1101
Db 386 ArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheSerSerVal 401
QY 1102 GAGACATCATGCAACACTACAGAAAT 1128
Db 402 ValGluLeuIleAsnHisTyrArgAsn 410

RESULT 5
A:38747
3-phosphatidylinositol 3-kinase (EC 2.7.1.-) 85k chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 16-Jul-1999
C:Accession: A38747
R:Escobedo, J.A.; Navanekasatun, S.; Kavanagh, W.M.; Milfey, D.; Fried, V.A.; Will
Cell 65, 75-82, 1991
A:Title: cDNA cloning of a novel 85 kd protein that has SH2 domains and regulates bin
A:Reference number: A38747; MUID:91191564; PMID:1849460
A:Accession: A38747
A:Molecule type: mRNA
A:Residues: 1-724 <ESC>
A:Cross-references: GB:M60651
C:Comment: This protein binds a phosphotyrosine-containing sequence of ligand-activat
3-phosphatidylinositol at position 3 of the inositol ring.
C:Superfamily: SH2 homology
C:Keywords: phosphotransferase
F:333-428/Domain: SH2 homology <SH2A>
F:624-718/Domain: SH2 homology <SH2>

```

Alignment Scores:

Pred. No.: 0.000654 Length: 724
 Score: 149.00 Matches: 76
 Percent Similarity: 32.76% Conservative: 39
 Best Local Similarity: 21.65% Mismatches: 96
 Query Match: 7.38% Indels: 140
 Gaps: 17

US-09-856-061-3 (1-1129) x A38747 (1-724)

```

OY 421 CCTTACACACCTCCGCGCTCTCATACACTCCGAGAGAGTACCAACCTTGCCCTC 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 85 ProthrProlyspProthrPro-----ProArg-----ProleuProVal 97
OY 481 GAGCGGAGAGCAGCAGG-----CCACT 504
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 98 AlaProGlySerSerLysThrGluAlaAspThrGluGlnGlnAlaLeuProleuProasp 117
OY 505 TTATCTCAGAGACACACCTTTCCAGAGAGTCCAGAGATGCCAGTCAAGCTTAAG 564
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 118 LeuAlaGluGlnPheAlaProProAspVal---AlaProProleuLeuLeuLeuLeu 136
OY 565 GACTTAAGTGAAGTC-----CTTGAAGCAGAAAAAGTTCTCATTAACCGAGAGCCT 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 137 GlnAlaIleGluLysGlyLysGlyLysSerThrLeuTyrArgThrGlnSerSer 156
OY 619 GAATCAACTCATCTGTAGAA-----AACCAAAATCTCAAGAG 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 157 AsnProAlaGluLeuArgGlnLeuLeuAspCysAspAlaAlaSerValAspLeuGluMet 176
OY 658 ATTCACATGGCCATTAAGAGTCTTCATTCACAGACAGACAGACAGTGCAGAAACAGA 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 IleAspValHisValLeuAlaAspAlaPheLysArgTyrLeuAlaAspLeuProasn 195
OY 718 GATCATAGAGAGGCGATGACGCCCTGTTCTCT 750
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 196 -----ProValIleProValAlaValTyrAsnGlnMetMetSer 208
OY 751 -----CAGAGATGCCAGCGCTCCAGCGCTGC 777
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 209 LeuAlaGlnGluLeuGlnSerProGluAspCysIleGlnLeuLeuLysLysLeuLeuArg 228
OY 778 -----AGCCCTCAGCAAAATATATCTGCCCTATTAATAC 810
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 LeuProAsnIleProHisGlnCysTrpLeuThrLeuGlnTyrLeuLeuLysHisPhePhe 248
OY 810 ----- 810
Db 249 LysLeuSerGlnAlaSerSerLysAsnLeuLeuAlaArgValLeuSerGluIlePhe 268
OY 810 ----- 810
Db 269 SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrGluHisLeuIleLys 288
OY 811 -----ACAAGCTGG 819
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 289 AlaIleGluIleLeuIleSerThrGluTrpAsnGluArgGlnProAlaProAlaLeuPro 308
OY 820 -----AAGCCACCTTCC 858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 309 ProLysProProLysProThrThrValAlaAsnAsnSerMetAspAsnAsnMetSerLeu 328
OY 859 CAGCAACATGAATGATACATGATGAGAAATACCGCGCAGCGAGTGAAGAGCATTCAG 918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 329 GlnAspAlaGluTrpTyrTrpLysAspIleSerArgGluGluValAsnGluLys---Leu 347
OY 919 AAGGAGAACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 348 ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrLysMetHisGlyAsp 367
OY 979 CCTTATGTTTGGCTGTGTTTATGAGAACAAAGTCTACAATGTAA---ATCCGCTTC 1035
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 368 ---TyrThrLeuThrLeu-----ArgLysGlyGlyAsnAsnLysLeuIleLysIle 383
  
```

OY 1036 CTGAGAGAGATCAGAGTGGCCCTGGGAGACAGCTCAGAGAGATGAGAGTTGAT 1095
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 384 PheHisArgAspGlyLysTyrGlyPheSerAspProLeu-----ThrPheAsn 399
 OY 1096 TCAGTAGAAGACATCATCGACACTCAGACATCAGAAAT 1128
 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
 Db 400 SerValValGluLeuLeuLeuAsnHisTyrArgAsn 410

RESULT 6

149553
 protein-tyrosine kinase (EC 2.7.1.112) emb - mouse
 N.Alternate names: B-cell progenitor cytoplasmic tyrosine kinase; Bruton agammaglobulin
 C.Species: Mus musculus (house mouse)
 C.Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Feb-2000
 C.Accession: I49553; B45184; J04771; PC4366
 R.Sideras, P.; Muller, S.; Shlets, H.; Jin, H.; Khan, W.N.; Nilsson, L.; Parkinson, E
 ulos, K.G.

J. Immunol. 153, 5607-5617, 1994
 A.Title: Genomic organization of mouse and human Bruton's agammaglobulinemia tyrosine
 A.Reference number: I49553; MUID:95081608; PMID:7989760

A.Accession: I49553
 A.Status: preliminary; translated from GB/EMBL/DBJ
 A.Molecule type: mRNA

A.Residues: 1-659 <SID>
 A.Cross-references: GB:L29788; NID:9625143; PIDN:AA66943.1; PID:9625144
 R.Tsukada, S.; Saffran, D.C.; Rawlings, D.J.; Parolimi, O.; Allen, R.C.; Klisak, I.;
 Cell 72, 279-290, 1993

A.Title: Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-link
 A.Reference number: AA5184; MUID:93145329; PMID:8425221
 A.Accession: B45184

A.Status: nucleic acid sequence not shown; not compared with conceptual translation
 A.Molecule type: mRNA

A.Residues: 1-122, P, 124-659 <TSU>
 A.Cross-references: GB:L08967; NID:9192233; PIDN:AA37316.1; PID:9192234
 A.Experimental source: 702/3 pre-B cell line

A.Note: sequence extracted from NCBI backbone (NCBIP:123834) and corrected to corresp
 R.Yamada, N.; Kawakami, Y.; Kimura, H.; Fukumachi, H.; Baier, G.; Altman, A.; Kato, T
 Biochem. Biophys. Res. Commun. 192, 231-240, 1993

A.Title: Structure and expression of novel protein-tyrosine kinases, Emb and Emt, in
 A.Reference number: J0471; MUID:93235578; PMID:8476425
 A.Accession: J0471

A.Status: nucleic acid sequence not shown
 A.Molecule type: mRNA

A.Residues: 1-66, A, 68-197, W, 198-449, A51-659 <YAM>
 A.Cross-references: GB:L10627; NID:9193016
 A.Note: this translation is not annotated in Genbank entry MUSEMBX, release 116.0

R.Kojima, T.; Fukuda, M.; Watanabe, Y.; Hamazato, F.; Mikoshiba, K.
 Biochem. Biophys. Res. Commun. 236, 333-339, 1997
 A.Title: Characterization of the pleckstrin homology domain of Btk as an inositol pol

A.Reference number: PC4366; MUID:97382431; PMID:9240435
 A.Accession: PC4366
 A.Molecule type: mRNA

A.Residues: 1-165 <KO>
 C.Comment: This protein specifically recognizes the inositol 1,3,4,5-pentakisphosphat
 e molecule to the membrane.

C.Genetics:
 A.Gene: emb; Btk
 C.Superfamily: protein-tyrosine kinase tec; pleckstrin repeat homology; protein kinas

C.Keywords: ATP; phosphoprotein; phosphotransferase; tyrosine-specific protein kinase
 F.2-131/Domain: pleckstrin repeat homology <PIK>
 F.221-269/Domain: SH3 homology <SH3>

F.281-377/Domain: SH2 homology <SH2>
 F.400-658/Domain: protein kinase homology <KIN>
 F.408-416/Region: protein kinase ATP-binding motif
 F.551/Binding site: phosphate (Tyr) (covalent) #status predicted

Alignment Scores:

Pred. No.: 0.00195 Length: 659
 Score: 143.00 Matches: 60
 Percent Similarity: 33.98% Conservative: 27
 Best Local Similarity: 23.44% Mismatches: 71
 Query Match: 7.08% Indels: 98

Db - 375 ProValProValAlaIalaProArgProProValGlnGlnIlyMetProMetProGlnAlaThr 394
 QY 817 TGGAGACCACCTTTCCCAAAAGGCTGATAGAAAGGATCCAGAC----- 864
 Db 395 ValArgProAlaValLeuProArgProGlnAsnThrProLeuProHisLeuGlnArgSer 414
 QY 864 ----- 864
 Db 415 ProProAspGlyInserPheArgGlyPheSerPheGluLysAlaArgInProSerGln 434
 QY 865 -----AATGAA 870
 Db 435 AlaAspThrGlyGluGluAspSerAspGlnAspTyrGluLysValProLeuProAsnSer 454
 QY 871 TGGTACATTGGAGATACAGCCCGCAGGACAGTGGAGAGAGCATTCATCAAG----- 921
 Db 455 ValPheValAsnThrThrInglInserCysGluValGlnArgLeuPheLysAlaThrAspPro 474
 QY 922 -----GAGAACACAGATGGTAGTTCTTGCTCCGAGATGTTCCACAAATCCAGGAA 975
 Db 475 ArgGlyGluProGlnAsnArgGlyLeuTyrCysIleArgAsnSerThrLysSerGly 494
 QY 976 GAGCCCATGTTGGCTGGCTTTATGAG-----AACCAAGCTCATCATGTAATAATC 1029
 Db 495 -----ValLeuValValTTPAspGlnSerSerLysValaArgAsnTyrArgIle 511
 QY 1030 CGCTTCTCGAGAGAAATCAGACAGTTGCGCTGGGAGCAGACAGATGAGAGAG 1089
 Db 512 -----PheGluLysAspSerLysPheTyr-----LeuGluGlyGluValLeu 525
 QY 1090 TTGTATTCAGTAGAAGACATCATCGAACTAC 1122
 Db 526 PheAlaSerValGlySerMetValGluHisTyr 536

RESULT 8
 A:5184
 protein-tyrosine kinase (EC 2.7.1.112), nonreceptor type, BTK - human
 N:Alternate names: atk; B cell progenitor kinase; Bruton agammaglobulinemia tyrosine kin
 C:Species: Homo sapiens (hmn)
 C:Date: 30-Apr-1993 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000
 C:Accession: I37212; I54541; I39457; S28912; A45184
 P:Ohta, Y.; Haire, R.N.; Litman, R.T.; Fu, S.M.; Nelson, R.P.; Kratz, J.; Kornfeld, S.J.
 Proc. Natl. Acad. Sci. U.S.A. 91, 9062-9066, 1994
 A:Title: Genomic organization and structure of Bruton agammaglobulinemia tyrosine kinase
 agammaglobulinemia.
 A:Reference number: I37212; MUID:94377492; PMID:8090769
 A:Accession: I37212
 A:Status: nucleic acid sequence not shown; translation not shown; not compared with conc
 A:Molecule type: DNA
 A:Residues: 1-659 <OHT>
 A:Cross-references: EMBL:U10087; NID:91226232; PIDN:AAB60639.1; PID:9517438
 A:Note: only Intron-exon junctions are shown
 R:Rohrer, J.; Parolino, O.; Belmont, J.W.; Conley, M.E.
 Immunogenetics 40, 319-324, 1994
 A:Title: The genomic structure of human BTK, the defective gene in X-linked agammaglobulin
 A:Reference number: I54541; MUID:95012452; PMID:7927535
 A:Accession: I54541
 A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBDUT
 A:Molecule type: DNA
 A:Residues: 1-659 <ROH>
 A:Cross-references: GB:I31572; NID:9632958; PIDN:AAA61479.1; PID:9632960
 R:Hagemann, T.L.; Chen, Y.; Rosen, F.S.; Kwan, S.P.
 Hum. Mol. Genet. 3, 1743-1749, 1994
 A:Title: Genomic organization of the Btk gene and exon scanning for mutations in patient
 A:Reference number: I39457; MUID:95152493; PMID:7860320
 A:Accession: I39457
 A:Status: Preliminary; translated from GB/EMBL/DBDUT
 A:Molecule type: DNA
 A:Residues: 1-278,281-659 <HAG>
 A:Cross-references: EMBL:U13433; NID:9575888; PIDN:AAC51347.1; PID:9575890
 R:Verla, D.; Vorchovsky, I.; Sideras, F.; Holland, J.; Davies, A.; Fliinter, F.; Hammar

A:Title: The gene involved in x-linked agammaglobulinemia is a member of the src family of protein tyrosine kinases

A:Reference number: S28912; MUID:93140868; PMID:8380905

A:Accession: S28912

A:Molecule type: mRNA

A:Residues: 1-659 <VERT>

R:Cross-references: GB:558957; NID:9312466; PIDN:CA44728.1; PID:9312467; GB:078027; R:Tsukada, S.; Saffran, D.C.; Rawlings, D.J.; Parolini, O.; Allen, R.C.; Kilsak, I.; Cell 72, 279-290, 1993

A:Title: Deficient expression of a B cell cytoplasmic tyrosine kinase in human x-link

A:Reference number: A45184; MUID:93145329; PMID:8425221

A:Accession: A45184

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-19, 'R', 21-190, 'A', 192-206, 'T', 208-210, 'T', 212-218, 'E', 220-238, 'E', 240-2

A:Cross-references: GDB:120542; OMIM:300300

A:Map position: Xq21.33-xq22

A:Introns: 47/3; 80/3; 103/3; 131/1; 174/1; 196/3; 259/2; 280/2; 298/3; 325/2; 368/1;

C:Superfamily: protein-tyrosine kinase tac; pleckstrin repeat homology; protein kinase

C:Keywords: ATP, phosphotransferase, tyrosine-specific protein kinase

F:2-131/Domain: pleckstrin repeat homology <PLK>

F:221-269/Domain: SH3 homology <SH3>

F:281-377/Domain: SH2 homology <SH2>

F:400-658/Domain: protein kinase homology <KIN>

F:408-416/Region: protein kinase ATP-binding motif

Alignment Scores:

Pred. No.:	0.00339	Length:	659
Score:	140.00	Matches:	58
Percent Similarity:	33.98%	Conservative:	29
Best Local Similarity:	22.66%	Mismatches:	71
Query Match:	6.93%	Indels:	98
DB:	2	Gaps:	11

US-09-856-061-3 (1-1129) x A45184 (1-659)

QY 373 AGAACCCAAACATTAAGAGAGATGCATCCGTAGAAAGACAGATTCCTTACACCT 432

Db 171 AtgagnglySerleuLysProGlySerSerHisArgLysThrLysLysProLeuProPro 190

QY 433 CCTCGCCTCTCATTAACACTTCGGAAGAAGTACCAACCCCTGGCCCGAGCCGAGAGC 492

Db 191 ThrProGluGluAspGlnIleLeuLysLys-----ProLeuProProGluProAlaIa 208

QY 493 AGCAGGCCACTTTATATCTCAGACACACCTTTCCAGAGTCCAGGAGTCCACATGCAG 552

Db 209 -----AlaPro 210

QY 553 ATTAACCTTAAGGAGCTTAAGTGAGGTCCTTGAGACA---GAAAAAGTTCCTCATTAACAG 609

Db 211 ValSerThrSerGluLeuLysValAlaIleuTyraSPtyrMetProMetAsnAla 230

QY 610 -----AGGAAGCCTGAATCAATCAATCAATCTATCTGTAGAAAACCAAAATACCTCAA 654

Db 231 AsnAspLeuGlnLeuArgLysGlyAspLysIlyrPheIleLeuGluGluSerAsn----- 248

QY 655 GAGATTCACACTTGCCATTAGACAGTTCTTCATTCACGACAGCAACACACAGTGTGCAAAAC 714

Db 249 ---LeuProTyr-----TyrArgAla 254

QY 715 AGAGATCATATAGAGAGGAGCATGACGCCCTGTTCTCTCAGAGATGCCAGCCCTCCAGCCAGC 774

Db 255 ArgAspLysAsnGly----- 259

QY 775 TGCACCCCTCAGCAAAATATATCGCCCTATATAATACACAGCTGGAGACACACTTCCCC 834

Db 260 -----GlnGluGlyTyrIleProSerAsnTyraIleThr----- 270

QY 835 AAAAGGTCTGATAGAAAGATGTCCAGCACATGAATGATGTGAGAAATACAGCCGC 894

Db 50 GluValThrHis---IleLysIleGlnAsnAsnGlyAspPheAspLeuTyGlyGly 68
 QY 679 TCCTTCATTCAGACACCAACACAGTGTGCAAAACAGACATCTATACAGAGCATCCAG 738
 Db 69 GluLysPheValThrLeuProGluLeuValGlnIleTyThrGluAsnGlyGluLeuLys 88
 QY 739 CCCTGTTCCTCTCAG-----AGATGCCACCTCCAGCCAGCTGACAGCCCTCAGCAAAAT 792
 Db 89 GluLysAsnGlnAlaIleGluLeuLysGlnProLeuIleCysAla----- 104
 QY 793 ATACTGCCCTATTAATACACAGCTGGAGACACACCTTCCCAAAAGCTGTGATAGAAAG 852
 Db 105 -----GluProThrThrGluArg----- 110
 QY 853 GATGTCCAGACATGAATGATCATGTGATGAGAAATACAGCCGCGGAGAGTGAAGAGCA 912
 Db 111 -----ThrPheAsnGlyAsnLeuSerGlyLysGluAlaGluLysLeu 124
 QY 913 TTCATGAGAGAGAACAGATGATGATGCTGTCGAGATTTGTCACAAATCCAG 972
 Db 125 IleuGluIleArgGlyLysAsnGlySerPheLeuValArgIle-----SerGlnSerLys 142
 QY 973 GAAGAGCCCATATGTTGGCTGTGTTTATGACAAAGTCAATGTAATAATCCGC 1032
 Db 143 ProGluAspPheValLeuSerValArgThrAspAspLysValThrHisValMetIleArg 162
 QY 1033 TTCCTGGAGAGATCGAGTTCCTGCTGCGGAGAGACTCAGAGAGATGAGAACTT 1092
 Db 163 Trp-----GlnAspLysLysTyThrAspValGly-----GlyGlyGluSerPhe 176
 QY 1093 GATTCACTAGACATCATCATGACACTACAG 1125
 Db 177 GlyThrLeuSerGluLeuIleAspHisLysTyLys 187
 RESULT 12
 T37550
 hypothetical coiled-coil protein [imported] - fission yeast (Schizosaccharomyces pombe)
 C:Species: Schizosaccharomyces pombe
 C:Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 08-Sep-2000
 C:Accession: T37550
 R:Conor, R.; Churcher, C.M.; McDougall, R.C.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, October 1999
 A:Reference number: Z21723
 A:Accession: T37550
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-485 <CON>
 A:Cross-references: EMBL:AL132714; PTDN:CA859797.1; GSPDB:GN00066; SPDB:SPAC11H1.01
 C:Experimental source: strain 972h-; cosmid c11H11
 C:Genetics:
 A:Gene: SPDB:SPAC11H1.01
 A:Map position: 1
 Alignment Scores:
 Pred. No.: 0.0157 Length: 485
 Score: 131.50 Matches: 74
 Percent Similarity: 40.41% Conservative: 65
 Best Local Similarity: 21.51% Mismatches: 118
 Query Match: 6.51% Indels: 87
 DB: 2 Gaps: 18
 US-09-856-061-3 (1-1129) x T37550 (1-485)
 QY 163 GAGCTTCGATGAGACATGCGACATGATTAATTTTACACAGCCGCGCTATGAAG 222
 Db 154 LysLeuLysValLysGluGluAsnGluAlaProProValIleProAlaLysProPheSer 173
 QY 223 GAATTCGATATATCAGATACACACTATTTCAAGTT---GCAATGCACT----- 270
 Db 174 SerSer-----SerGluGlnHisPheArgLysValProAlaLeuProSerLysLeuPro 191
 QY 271 CCCCTTCGCTTACAGACAGACGCTCTATGTCATTCGATGGAAGCCGAGCTTGAACACACAG 330
 Db 271 CCGCTTCGCTTACAGACAGACGCTCTATGTCATTCGATGGAAGCCGAGCTTGAACACACAG 330

Db 192 ProLysProLeuLysIleThrAlaAsnSerLeuGlyGlnGluThrAsnSerAsnSer 211
 QY 331 ACGAGGTTGGAAAGAGG-----GACAAACCATTTCCAGGAGACTCAGA 375
 Db 212 SerSerPheGlnSerThrLeuPheSerLeuAsnThrAlaProPheSer-----AlaThr 229
 QY 376 AGCCAAACATTAAGAGATGCA---TCCGTAAAGAAAGACAGATTCCTTACCACT 432
 Db 230 SerGlnGlnLeuValHisAspSerValSerLeuArg----- 241
 QY 433 CTTGCGCTCTCATACACTTCCAGAGAGTACCAACCTTGGCCCTGAGCCGAGAGC 492
 Db 242 ---ArgProSerSerAsnIleProAlaGln---LysProIleProProLysProGluGln 259
 QY 493 AGCAGG-----CCACCTTATCTCGAAGACACACCTTCCAGAA 531
 Db 260 AsnIleIleIleIleThrLysAspThrProSerLeuLysAspLysTySerLysProAla 279
 QY 532 GTCCAGGGAATGCCAGTACATGACCTTAAGGACTTAAGTGAAGCTTGAAGAGAA 591
 Db 280 LeuLeuProGlnLysProLysValSer-----LysGlyGlnIleValGlnGlnVal 296
 QY 592 AAATTCCTCATTAACAGAGAGAGCTGAAATCAATCATCATCTGTTGAAGAAACAAATACT 651
 Db 297 SerValPheSerThrGlyLysLysIleGluSerGlnSerLeuLeuAsnLeuIleAspThr 316
 QY 652 CAA---GAGATTCACCTTCGCATTAAGCATTCCTTCATTCACGAAACCAACCAAGTGTG 708
 Db 317 AspIleGluThrProLeuLysGlySer-----GluLeuLeu 329
 QY 709 CAAGACAGATCATAG 768
 Db 330 TyrSerGluAspPheLysProAsnValAspProValLysIleGlnGlnIle----- 346
 QY 769 GCCAGCTGCAGCCCTCAGAAATATGCTGCCCTATTAATACAACTGTG----- 819
 Db 347 -----LeuHisLysGlnAsnLysIleIleGluLysLysPheSerGlnIleArgIle 364
 QY 820 -----AGACCACTTCCCAAAAGCTC 843
 Db 365 SerLysAsnLeuGluValLysGlnArgLeuLeuAspGlnGluArgHisAlaLeuGluThr 384
 QY 844 GATGAAAGATGTCCAGACACATGAATGATGATGAGAGATACAGCCGCGAGAGAGTGTG 903
 Db 385 LeuAlaLysAsnIleGluAsnAsnArgPheIleLeuGlyLysArgValGlyAla--- 403
 QY 904 GAAGAGCATTCATGAG 963
 Db 404 ArgGluAlaLeuGlnLysLeuAspAsn-----LeuLysAspLeuSerVal 418
 QY 964 AAATCCAAAG 1023
 Db 419 -----GlnGluLeuPheIleIleIleProSerGluArgGluLeuLysTyThrGluLeu 435
 QY 1024 AAATTCGCTTCTCGAGAGAGATCAGACTTGCCTGGGAGACAGAGACTCAGAGAGAT 1083
 Db 436 Lys-----ArgLysAsp 439
 QY 1084 GAGAGATTGAT 1095
 Db 440 GluLysLeuAsp 443
 RESULT 13
 B38749
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85k chain B - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 16-Jul-1999
 C:Accession: B38749
 R:Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, J.
 Cell 65, 91-104, 1991
 A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine
 A:Reference number: A38749; M0ID:91191567; PMID:1707345

A:Accession: B38749
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-723 <OTS>
 A:Cross-references: GB:M61745; GB:M61746
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:325-420/Domain: SH2 homology <SH2A>
 F:617-706/Domain: SH2 homology <SH2>

Alignment Scores:

Pred. No.:	0.018	Length:	723
Score:	131.00	Matches:	81
Percent Similarity:	34.75%	Conservative:	50
Best Local Similarity:	21.49%	Mismatches:	118
Query Match:	6.49%	Indels:	128
	2	Gaps:	18

US-09-856-061-3 (1-1129) x B38749 (1-723)

```

QY 301 TCATTTGGA---CAGCCGACCTGGACACACAGACGAGTGGAGAGAGTGCACAAACC 357
    |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 53 SerValGlyTyrMetProGlyLeuAsnGluArgThrArg---GlnArgGlyAspPhePro 71
QY 358 ATTTCCAGGAGCGTCGAAAGCCAAACATTAAAGAGATGCATCCCTTAAGAAAGAACAG 417
    |||:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 72 GlyThr-----TyrValGluPheLeuGlyProValAlaLeuLeuAlaArg----- 85
QY 418 ATTCCCTTACGACCTCTCGGCTCTCATACACTTCCGAGAGATACCAACCCCTTGCCC 477
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 86 ---ProGlyProArgProArgProGlyProArgProLeuProAlaArgProArgAspGlyPro 104
QY 478 CCTGAGCCGAGAGCAGCAGCCACCTTATCTCAGACACACCTTCCAGAGTTCAG 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 105 ProGluProGlyLeuThrLeuProAspLeuProGluGlnPheSerProProAspVal--- 123
QY 538 GGAATGCCAGTCAGATTAAGCTTAAGGAGCTCTTGAAGCAGAAAGATT 597
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 AlaProProIleLeuValAlaLeu-----ValGluAlaIleGluArgThr 138
QY 598 CCTCATACACAGAGAGCCTGAAATCACTCATCTGTAGAAACCAAAATACTCAGAG 657
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 139 GlyLeuAspSerTyrArgProGluProAlaValArgThrAspPheSerLeuSerAsp 158
QY 658 ATTTCCA-----CTTGCATTT 672
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 159 ValGluGlnTTPAspAlaAlaAlaLeuSerAspGlyValIleGlyPheLeuLeuAlaLeu 178
QY 673 AGCAGTTCTTCATTCACGACAAAC----- 696
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 179 ProAlaProLeuValThrProGluAlaAlaAlaGluAlaHisArgAlaLeuArgGluAla 198
QY 697 -----AACCAAGTGTGCAAAAC 714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 199 AlaGlyValGlyProAlaLeuGluProProThrLeuProLeuHisAlaLeuThrLeu 218
QY 715 AGA-----GATCATAGAGAGGAGCAGACCCCTGTTCTCCCTAG----- 753
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 ArgPheLeuLeuGlnHisLeuGlyArgValAlaGlyArgAlaProAlaProGlyProAla 238
QY 754 -----AGATGCCAGCCTCCAGCAGCAG 774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 ValArgAlaLeuGlyAlaThrPheGlyProLeuLeuLeuArgAlaProProProser 258
QY 775 TGCAGCCCT----- 783
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 259 ProProProGlyGlyAlaProAspGlyThrGluProThrProAspPheProAlaLeuLeu 278
QY 784 -----CACGAATAATAGTCCCTATAATAACACA 813
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 279 ValGlyLeuLeuLeuGlnHisLeuGluGluGlnValAlaProProAlaLeu----- 297
QY 814 AGCTGGAGACACCTTTCCCAAAAGTCT----- 843

```

```

Db 298 -----ProProLysProProLysThrLysProAlaProThrGlyLeuAlaAsnGly 314
QY 844 ----GATAGAAAGATCTCCAGACCAATGATGATGATGATGAGATATACGCCGCCAGCA 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 GlySerProProSerLeuAsnAlaGluIlePtyrTyrPheGlyAspIleSerArgGluGlu 334
QY 901 GTGAGAGAGCATTCATGAGAGGAGAACAGATGATGATGATGATGATGATGATGATGATG 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 335 ValAsnGluLys---LeuArgAspThrProAspGlyThrPheLeuValArgAspAlaSer 353
QY 961 ACAAAATCCAGAGAGAGAGCCCTATGTTTGGCTGTGTTTATGAGAAACATTCATCAAT 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 SerLysIleGlnGlyLys---TyrThrLeuThrLeu-----ArgLysGlyGlyAsn 369
QY 1021 GTTAAAT---ATCGCTCTCTCGAGAGAGATTCAGCACTTTGCCCTGGGAGACCTCACA 1077
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 370 AsnLysLeuIleLysValPheHisArgAspGlyHisTyrGlyPheSerGluProLeu--- 388
QY 1078 GGAGATGAGAGATTGATTGATTCAGTGAAGACATCATCATGACACTGACAGAT 1128
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 389 -----ThrPheCysSerValValAspLeuIleThrHisTyrArgHis 402

```

RESULT 14
 H59435
 phosphoinositide-3-kinase regulatory beta chain [imported] - huamm
 C:Species: Homo sapiens (man)
 C:Date: 03-Jun-2002 #Sequence-revision 03-Jun-2002 #text-change 03-Jun-2002
 C:Accession: H59435; A59436
 R:Volin S; Patrachini P; Otsu M; Hiles I; Gout I; Calzolari E; Bernardi F; Rooke L
 Oncogene 7, 789-793, 1992
 A:Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinosit
 A:Reference number: H59435
 A:Accession: H59435
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <VOL>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1
 R:Janssen, J.W.; Schleithoff, L.; Barttram, C.R.; Schulz, A.S.
 Oncogene 16, 1767-1772, 1998
 A:Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta sub
 A:Reference number: A59436; MUID:98241181; PMID:9582025
 A:Accession: A59436
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <GAN>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1

Alignment Scores:

Pred. No.:	0.0413	Length:	728
Score:	126.50	Matches:	75
Percent Similarity:	34.34%	Conservative:	27
Best Local Similarity:	25.25%	Mismatches:	108
Query Match:	6.26%	Indels:	88
	2	Gaps:	15

US-09-856-061-3 (1-1129) x H59435 (1-728)

```

QY 287 CCAGAGCTCTA---TTCCTATTGGACAGCCGACCTGGAACACACAGACGAGTGGAA 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 ProAlaProLeuValThrProGluAlaSerAlaGluAla-----ArgArgAlaLeuArg 199
QY 344 GAGTGGACAACCACTTCCAGGAGAGTCAGAAAGCCAAACATTAAAGAGATGATCCG 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 200 GluAlaAlaGlyProValGlyProAlaLeuGluProProThrLeu---ProLeuHisArg 218
QY 404 TAAGAAAGAACAGATTCCTTACACACTCTCGGCT----- 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 219 AlaLeuThrLeuArgPheLeuLeuGlnHisLeuGlyArgValAlaArgArgAlaProAla 238
QY 442 -----CTCATACACTTCGAGAGAGTACCAACCTTG-----CCGCTT 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 239 LeuGlyProAlaValArgAlaLeuGlyAlaThrPheGlyProLeuLeuLeuArgAlaPro 258

```

```

QY 481 GAGCCGAGACGACGACCTTTATCTCAGAGACACACCTTTCCAGAGTCCAGGA 540
    ||| ||||| ||||| |||
Db 259 ProProSerSerProProPro-----GlyGly 268
QY 541 ATGCCAGTCAGATTAAGCTTAAGGACTTAAGTGCCTTGAAGAGAAAAAGTTCT 600
    |||
Db 269 AlaProAsp----- 271
QY 601 CATAGCAGAGAGACCTGATCACTATCTGTAGAAAACCAAAATACTCAAGACTT 660
    : : : : : ||||| : : :
Db 272 GlySerGluProSerProAspPheProAlaLeuLeuValGluLysLeuGluGlu--- 290
QY 661 CCAGCTTCATTAAGCAGTTCTTCATTCAGCAGACAAACACAGTGTCAAAACAGAGAT 720
    : : : : :
Db 290 ----- 290
QY 721 CATAGAGAGAGCAGACGCTGTTCTTCAGAGATGCCAGCTCCAGCAGCTGCAGC 780
    : : : : : |||
Db 291 HisLeuGluGluGluGluValAlaProProAlaLeuProProLysProProLysAlaLys 310
QY 781 CCTCAGCAAAATATACGCTCCTATTAATACACAGCTGAGACACCACTTTCCAAAAAG 840
    ||| : : : : : |||||
Db 311 ProAlaProThrValLeuAla-----AsnGlyGlySerProPro----- 323
QY 841 TCTGATAGAAAGATGTCACGACACATGATGATGATGAGAGATACGCCGACGAGCA 900
    : : : : : ||||| |||||
Db 324 -----SerLeuGlnAspAlaGluLysPheLysPheLysPheLysPheLysPheLys 339
QY 901 GTGCAAGAGCATTCAATGAGAGAGACACAGATGATGATGATGATGATGATGATGATG 960
    ||| ||| : : : : : ||||| |||||
Db 340 ValAsnGluLys---LeuAlaArgAspThrProAspGlyThrPheLeuValAlaArgAspAlaSer 358
QY 961 ACAAAATCCAGAGAGACCCCTATGTTTGGCTGGTGTATGAGAACAAAGTCTACAT 1020
    : : : : : ||| : : : ||| : : : |||
Db 359 SerLysIleGluGluGlu---TyrThrLeuThrLeu-----ArgLysGlyGlyAsn 374
QY 1021 GTAAA---ATCGCTTCCTCGAGAGATCAGCAGTTTGCCCTGGGACAGACTCAGA 1077
    ||| : : : : : ||||| : : : |||
Db 375 AsnLysLeuIleLysValAlaPheHisArgAspGlyHisTyrGlyPheSerGluProLeu--- 393
QY 1078 GGAGATGAGAGATTGATTCAGTAGAGACATCATCGACACTACAGAAAT 1128
    ||| ||||| ||||| |||||
Db 394 -----ThrPheCysSerValValAspLeuIleAsnHisTyrArgHis 407

```

RESULT 15

S08636

nck protein - human

N/Alternate names: src-related protein

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 05-Nov-1999

C/Accession: S08636

R/Lehmann, J.M.; Rietmuller, G.; Johnson, J.P.

Nucleic Acids Res. 18, 1048, 1990

A/Title: Nck, a melanoma cDNA encoding a cytoplasmic protein consisting of the src homol

A/Reference number: S08636; MUID:90192089; PMID:2107526

A/Accession: S08636

A/Molecule type: mRNA

A/Residues: 1-377 <LEH>

A/Cross-references: EMBL:X17576; NID:g35014; PIDN:CAA35599.1; PID:g35015

C/Superfamily: SH3 homology; SH2 homology

F/9-56/Domain: SH3 homology <SH31>

F/113-160/Domain: SH3 homology <SH32>

F/197-247/Domain: SH2 homology <SH33>

F/282-371/Domain: SH2 homology <SH2>

Alignment Scores:

```

Pred. No.: 0.0504 Length: 377
Score: 125.00 Matches: 33
Percent Similarity: 40.00% Conservative: 23
Best Local Similarity: 23.57% Mismatches: 52
Query Match: 6.19% Indels: 32
DB: 2 Gaps: 4

```

```

US-09-856-061-3 (1-1129) x S08636 (1-377)
QY 706 GTGCAAAACAGAGATCATAGAGAGCAGTGCAGCCCTGTTCTCTCAGAGATGCCAGCT 765
    : : : : : ||||| : : : : : |||||
Db 250 MetGlnAsnAsnProLeuThrSerGlyLeuGluProSerProProGln----- 265
QY 766 CCAGCAGCTCAGACCCCTCAGAAATATTAAGTGGCCCTTAATATCACAGCTGGAGACA 825
    |||
Db 266 -----CysAspTyrIleArgProSerLeuThrGlyLysPheAlaGly----- 279
QY 826 CTTTCCCAAAAGTCTGATAGAAAGATGTCAGCAGCAATGATGATGATGATGATGATG 885
    : : : : : ||| : : : : : |||||
Db 280 -----AsnProTyrTyrGlyLys 286
QY 886 TACAGCCGCGCAGCAGTGAAGAGGATTCATGAAGAGAAACAAGATGATGATGATG 945
    : : : : : ||| : : : : : |||||
Db 287 ValThrArgHisGlnAlaGluMetAlaLeuAsnGluArgGlyHisGluLysAspPheLeu 306
QY 946 GTCCGAGATTGTCACAAATCCAGAGAGAGCCCTATGTTTGGCTGTGTTTATGAG 1005
    : : : : : ||| : : : : : |||||
Db 307 IleArgAspSerGluSerSerProAsnAspPheSerValSerLeuLysAlaGluGlyLys 326
QY 1006 AACAAAGTCTCAATGTAAAAATCCGCTTCCTGAGAGAGAAATCAGCACTTGCCCTGGG 1065
    ||||| : : : : : ||||| : : : : : |||||
Db 327 AsnLysHisPheLysValGlnLeu-----LysGluThrValTyrCysIle--- 341
QY 1066 ACAGAGCTCAGAGAGATGAGAAAGTTCATTCAGTAGAAGACATCATCGACACTACAG 1125
    ||| ||||| ||||| : : : : : |||||
Db 342 -----GlyGlnArgLysPheSerThrMetGluGluLeuValGlnHisTyrLys 357

```

Search completed: April 21, 2003, 12:32:13
 Job time : 48.2565 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: April 21, 2003, 12:15:11 ; Search time 16.2418 seconds
(without alignments)
5766.221 Million cell updates/sec

Title: US-09-856-061-3
Perfect score: 2020
Sequence: 1 ttcacgaactcagctgcgc.....catcgacactacaagaatt 1129

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 225784

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame-n2p.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09856061/rnrat.21042003_113018_3735/app_query.fasta.1.3150
-DB=SwissProt_40 -QWRT=fastan -SWFRT=n2p.rsp -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORMEXT -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09856061 @CGN.1.1.65 @rnrat.21042003_113018_3735 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEROUT=120
-WARN_TIMEROUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332.5	16.5	533	1	LCP2_HUMAN
2	315.5	15.6	533	1	LCP2_MOUSE
3	154.5	7.6	724	1	P85A_BOVIN
4	150.5	7.5	724	1	P85A_HUMAN
5	149	7.4	724	1	P85A_MOUSE
6	145	7.2	724	1	P85A_RAT
7	143	7.1	659	1	BTX_MOUSE
8	140	6.9	559	1	3BP2_MOUSE
9	136.5	6.9	659	1	BTX_HUMAN
10	136.5	6.8	645	1	CSW_DROME
11	136	6.7	561	1	3BP2_HUMAN
12	133.5	6.6	724	1	P85B_BOVIN
13	128	6.3	461	1	P55G_HUMAN
14	128	6.3	722	1	P85B_MOUSE
15	127	6.3	271	1	CRK_DROME
16	126.5	6.3	533	1	SOC6_MOUSE
17	126.5	6.3	728	1	P85B_HUMAN
18	126.5	6.3	1490	1	CRK7_HUMAN

19	126	6.2	461	1	P55G_BOVIN	046404 bos taurus
20	125	6.2	377	1	NCK1_HUMAN	P16333 homo sapien
21	125	6.2	585	1	PTNB_MOUSE	P35235 mus musculu
22	125	6.2	593	1	PTNB_HUMAN	006124 homo sapien
23	125	6.2	593	1	PTNB_RAT	P41499 ratius norv
24	124.5	6.2	217	1	GRAP_HUMAN	013588 homo sapien
25	123.5	6.1	722	1	P85B_RAT	063788 ratius norv
26	123.5	6.1	1196	1	ABLI_CAEEL	P03949 caenorhabd
27	123	6.1	461	1	P55G_MOUSE	064143 mus musculu
28	122	6.0	822	1	PER_HUMAN	P16591 homo sapien
29	120.5	6.0	833	1	YMY2_CAEEL	P34531 caenorhabd
30	120	5.9	847	1	VAV3_HUMAN	09ukw4 homo sapien
31	120	5.9	939	1	ST20_YEAST	003497 saccharomyc
32	119	5.9	620	1	ITK_HUMAN	008881 homo sapien
33	119	5.9	1872	1	T2D1_HUMAN	P21675 homo sapien
34	116	5.7	517	1	FCR_MOUSE	P14234 mus musculu
35	115	5.7	675	1	BMX_HUMAN	P51813 homo sapien
36	115	5.7	1290	1	PIG1_HUMAN	P19174 homo sapien
37	114.5	5.7	625	1	ITK_MOUSE	003526 mus musculu
38	114.5	5.7	630	1	TEC_MOUSE	P24604 mus musculu
39	114	5.6	1128	1	BEM3_YEAST	P32873 saccharomyc
40	114	5.6	1190	1	ZO2_HUMAN	09udg2 homo sapien
41	114	5.6	2440	1	NCK1_HUMAN	075376 homo sapien
42	112.5	5.6	520	1	YEAB_SCHPO	014079 schizosacch
43	112.5	5.6	303	1	CRKL_HUMAN	P46109 homo sapien
44	112.5	5.6	878	1	VAV2_HUMAN	P52735 homo sapien
45	112.5	5.6	1942	1	Y054_HUMAN	P42694 homo sapien

ALIGNMENTS

RESULT 1

LCP2_HUMAN STANDARD; PRT; 533 AA.

AC Q13054; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).

GN LCP2.
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;
RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE=Leukemia;
RX MEDLINE=95221345; PubMed=7706237;

RA Jackman J.K., Motto D.G., Sun Q., Tanemoto M., Turck C.W., Peltz G.A.,
RT Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein
associated with Grb2 in T cells.;

RL J. Biol. Chem. 270:7029-7032(1995).

RN [2]
RP SEQUENCE FROM N.A.

RC TISSUE=Prostate;
RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.

CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND FYB.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).

CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND
PERIPHERAL BLOOD LEUCOCYTES. HIGHLY EXPRESSED ALSO IN T CELL AND
MONOCYTIC CELL LINES. EXPRESSED AT LOWER LEVEL IN B CELL LINES.

CC NOT DETECTED IN FIBROBLAST OR NEUROBLASTOMA CELL LINES.

CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION BY ZAP-70.

CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -1- DATABASE: NAME=PROT; NOTE=PROT 1.1:-5(2000);
WWW="http://www.ncbi.nlm.nih.gov/prot/guide/1118450040_g.htm".

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation.

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
 DR EMBL: U20158; AAC50135.1; -
 DR EMBL: BC016618; AAH16618.1; -
 DR HSSP: P12931; 1SHD.
 DR Genew: HGNC:6529; LCP2.
 DR MIM: 601603; -
 DR InterPro: IPR001660; SAM.
 DR InterPro: IPR000980; SH2.
 DR Pfam: PF00017; SH2; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00454; SAM; 1.
 DR SMART: SM00252; SH2; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR SH2 domain; Phosphorylation.
 FT DOMAIN 133 136 POLY-GLU.
 FT DOMAIN 198 201 POLY-PRO.
 FT DOMAIN 422 530 SH2.
 SQ SEQUENCE 533 AA; 60188 MW; C5D22F31D36200C8 CRC64;

Alignment Scores:
 Pred. No.: 1.09e-17 Length: 533
 Score: 332.50 Matches: 113
 Percent Similarity: 36.24% Conservative: 49
 Best Local Similarity: 25.28% Mismatches: 128
 Query Match: 16.46% Indels: 157
 DB: 1 Gaps: 16

US-09-856-061-3 (1-1129) x LCP2_HUMAN (1-533)

OY 58 GGCCAGTACCAAGATGACAAAGCCTCTTACAGTGGGAAAGAACTTCTGCACGTC 117
 DB 126 GYAAsPTyrgLusSerProasn-----GluGluGluGluAlaProVal 139
 OY 118 CTGATGTGAGCAAAAGCCACATGATGATGATGATGATGATGATGATGATGATGAT 177
 DB 140 GUAAsP-----AspAlaAspTyrgLusProProProSerSerSerSerSer 153
 OY 178 GAGACATGCGACGTGATTAATAATTTTACACGCGCGCTATTAAGAAATGTAATATGCA 237
 DB 134 GUAAlaLeuGlnAsn---SerIleLeuProAlaLysProPheProAsnSer----- 169
 OY 228 GATTAACACTATTTCAGAGTTGCATGACACTCCCTCCCTTACAGACACGACCTCT 297
 DB 170 AsnSerMetTyrr----- 173
 OY 238 ATCTCATTTGGACAGCCGACCTGGAACACACAGAGAGTTGGAAGAAGTGACAAACC 357
 DB 174 ----- 177
 OY 358 ATTTCAGAGGAGCTCAGAAAGCCAAACATTAAGAGATGATCGCTAAGAAAGAACAG 417
 DB 178 ProSerGlyLysThrProGlnGlnPro----- 186
 OY 418 ATTCTTTACCACTCCCTCGGCTCTCATACACTTCCG-----AAGAAG 462
 DB 187 ---ProValProProGlnThrProMetAlaAlaLeuProProProProAlaGlyArgAsn 205
 OY 463 TACCAACCTTGCCCTC-----GAGCCGAGAGAGCAGAG----- 498
 DB 206 HisSerProLeuProProProGlnThrAsnHisGluGluProSerArgSerArgSerHis 225
 OY 498 ----- 498
 DB 226 LysThrAlaLysLeuProAlaProSerIleAspArgSerThrLysProProLeuAspArg 245
 OY 499 -----CCACTTTATCTCAG 513
 DB 246 SerIleAlaProPheAspArgGluProPheThrLeuGlyLysLysLysProProPheSerAsp 265

OY 514 AGACACACCTTTCGA-----GAAGTCCAG 537
 DB 266 LysProSerIleProAlaGlyArgSerIleuGlyGlnHisLeuProLysIleGlnLysPro 285
 OY 538 GGAATGCCCACTCAGATAAGCTTAAGGACTTAAGTGAAGTCTTGAAGCAGAAAGTT 597
 DB 286 ProLeuProProThrThrGlnArgHisGluArgSerSerProLeuProGlyLysLysPro 305
 OY 598 CCT-----CAT 603
 DB 306 ProValProLysHisGlyTrpGlyProAspArgGlyAsnAspGluAspAspValHis 325
 OY 604 AACCAAGAGAGAGCTGATCACTCATCTGTAGAAACCAAAATACT----- 651
 DB 326 GlnArgProLeuProGlnProAlaLeuLeuProMetSerSerSerAsnThrPheProSerArg 345
 OY 652 -----CAAGAGATTCCACTTCCCATTAAGCAGTTCTTCATTCACG 690
 DB 346 SerThrLysProSerProMetAsnProLeuProProSerSerHisMetProGlyAlaPheSer 365
 OY 691 ACAAGCAACCAAGTGTGCAAAACAGAGATCATAGAGAGCATGACGCCCTGTCTCTCT 750
 DB 366 GluSerAsnSerSerPhe-----ProGlnSerAlaSerLeuProProTyrrPheSer 382
 OY 751 CAG-----AGATGCCAGCCTCCAGCCAGCTGACGCCCTCAGCAAAATATATCTGCCCTAT 804
 DB 383 GlnGlyProSerAsnArgProProIleArgAlaGluGlyArgAsnPheProLeuProLeu 402
 OY 805 AAATACACAAGCTGGAGACACACCTTCCCAAAAGCTGTGATGAGAAGATGTCAGACAC 864
 DB 403 Pro---AsnLysProArgProProSerPro-----AlaGluGluGluAsnSerLeuAsn 419
 OY 865 AATGATGCTATATGATGAGAAATACAGCCGCCAGCATGCGAAGAGCATTCATGAAAGCG 924
 DB 420 GlnGluTrpTyrrValSerTyrrIleThrArgProGluAlaGluAlaLeuArgLysIle 439
 OY 925 AACCAAGATGAGTATTTCTTGTGCGAGATTTTCCACAAATCCAAAGAGAGCCCTAT 984
 DB 440 AsnGlnAspGlyThrPheLeuValArgAspSerSerLysLysThrThrThrAsnProTyrr 459
 OY 985 GTTTGGCTGTGTTTATGAGAACAAAGTCTAATGTMAAATCCGCTTCTCGAGAGG 1044
 DB 460 ValLeuMetValLeuLeuGlyLysAspLysValTyrrAsnIleGlnIleArgTyrrGlnLysGlu 479
 OY 1045 AATCAGCAAGTTGCCCTGGGAGCAGCATCAGAGAGATGAGAAATTTGATTCATGAGA 1104
 DB 480 SerGlnValTyrrLeuLeuGlyThrGlyLeuArgGlyLysGluAspPheLeuSerValSer 499
 OY 1105 GACATCATCGAAGACTACAG 1125
 DB 500 AspIleIleAspTyrrPheArg 506

RESULT 2
 LCP2_MOUSE STANDARD: PRT: 533 AA.
 AC 060787;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte
 DE protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).
 GN LCP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N. A.
 RC TISSUE=T-cell lymphoma;
 RX MEDLINE=95221345; PubMed=7706237;
 RA Jackson J.K., Motlo D.G., Sun Q., Tanemoto M., Turck C.W., Pelicz G.A.,
 RA Koretzky G.A., Finkel P.R.;

RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Larrea F., Panayotou G.,
 RA Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,
 RA Courtneidge S.A., Parker P.J., Waterfield M.D.;
 RT "Characterization of two 85 kd proteins that associate with receptor
 RT tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase.";
 RL Cell 65:91-104(1991).
 [12]
 RP CIRCULAR DICHROISM AND FLUORESCENCE SPECTROSCOPY.
 RA MEDLINE=93049176; PubMed=1330535;
 RA Panayotou G., Bax B., Gout I., Federwisch M., Wrzobowski B., Dhand R.,
 RA Fry M.J., Blundell T.L., Moliner A., Waterfield M.D.;
 RT "Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
 RT domain with a PDGF receptor phosphorylation site: structural features
 RT and analysis of conformational changes.";
 RL EMBO J. 11:4261-4272(1992).
 [13]
 RP STRUCTURE BY NMR OF 1-84.
 RA MEDLINE=93272320; PubMed=7684655;
 RA Bocker G.W., Gout I., Downing A.K., Driscoll P.C., Boyd J.,
 RA Waterfield M.D., Campbell I.D.;
 RT "Solution structure and ligand-binding site of the SH3 domain of the
 RT p85 alpha subunit of phosphatidylinositol 3-kinase.";
 RL Cell 73:813-822(1993).
 [14]
 RP STRUCTURE BY NMR OF 314-431.
 RA MEDLINE=93257146; PubMed=1323062;
 RA Bocker G.W., Breeze A.L., Downing A.K., Panayotou G., Gout I.,
 RA Waterfield M.D., Campbell I.D.;
 RT "Structure of an SH2 domain of the p85 alpha subunit of
 RT phosphatidylinositol-3-OH kinase.";
 RL Nature 358:684-687(1992).
 [15]
 RP STRUCTURE BY NMR OF 321-434.
 RA MEDLINE=97110350; PubMed=8952511;
 RA Guenther U.L., Liu Y., Sanford D., Bachovchin W.W., Schaffhausen B.;
 RT "NMR analysis of interactions of a phosphatidylinositol 3'-kinase SH2
 RT domain with phosphotyrosine peptides reveals interdependence of major
 RT binding sites.";
 RL Biochemistry 35:15570-15581(1996).
 [16]
 RP STRUCTURE BY NMR OF 614-724.
 RA MEDLINE=96173872; PubMed=9512716;
 RA Siegal G., Davis B., Kristensen S.M., Sankar A., Linares J.,
 RA Stein R.C., Panayotou G., Waterfield M.D., Driscoll P.C.;
 RT "Solution structure of the C-terminal SH2 domain of the p85 alpha
 RT regulatory subunit of phosphoinositide 3-kinase";
 RL J. Mol. Biol. 276:461-478(1998).
 [17]
 RP FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC [1] SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNTS.
 CC [1] SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC [1] SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC [1] SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M61745; AAA79511.1; -
 DR PIR: A38749; A38749.
 DR PDB: 2PNA; 31-JAN-94.
 DR PDB: 2PNB; 31-JAN-94.
 DR PDB: 2PNI; 31-OCT-93.
 DR PDB: 1PNJ; 31-OCT-93.
 DR PDB: 1BFI; 25-FEB-98.

DR PDB: 1BFI; 25-FEB-98.
 DR InterPro: IPR001720; PI3kinase_P85.
 DR InterPro: IPR000198; RHOGAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RHOGAP; 1.
 DR PRINTS: PR00678; PI3KINASEP85.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRODOM: PD000066; SH3; 1.
 DR PRODOM: PD000093; SH2; 2.
 DR SMART: SM00324; RHOGAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 KW SH3 domain; SH2 domain; Repeat: 3d-structure.
 FT DOMAIN 3 79
 FT DOMAIN 129 273
 FT DOMAIN 333 428
 FT DOMAIN 624 718
 FT HELIX 340 347
 FT TURN 348 348
 FT STRAND 354 359
 FT STRAND 368 373
 FT STRAND 378 382
 FT STRAND 384 384
 FT STRAND 391 391
 FT STRAND 391 391
 FT HELIX 401 409
 FT STRAND 413 413
 FT TURN 418 418
 FT STRAND 422 422
 FT STRAND 427 427
 SQ SEQUENCE 724 AA; 83497 MW; EBD6E6754BBF7321 CRC64;
 Alignment Scores:
 Score: 0.000324 Length: 724
 Percent Similarity: 154.50 Matches: 81
 Best Local Similarity: 37.21% Conservative: 47
 Query Match: 23.5% Mismatches: 102
 DB: 7.65% Indels: 114
 Caps: 14
 US-09-856-061-3 (1-1129) x P85A_BOVIN (1-724)
 QY 406 AGAAGAACAAAGATCTCTTACCACTCTCGGCTCTCATTAACACTTCCGACAGATAC 465
 Db :::: 111 111 111111111 111111
 Db 80 LysLysLysSerProProThrProLysProArgProProArgProLeuProValaLaPro 99
 QY 466 CAACCTTGCCCGCCGACCGGAGAGACGAGG-----CCACCTTATCT 510
 Db :::: 111 111 :::: ::::
 Db 100 GlyProSerLysThrGlnAlaAspSerGlnGlnAlaSerThrLeuProAspLeuAla 119
 QY 511 CAGAGACACACCTTCCGAGAGTC----- 534
 Db :::: 111111111 :::: 111
 Db 120 GluGlnPheAlaProProAspValaLaProProLeuLeuIleLysLeuValGlnValaLe 139
 QY 535 -----CAGGGAATGCCCAAGTCAGATAGCTTAAG----- 564
 Db :::: 111111111 :::: 111
 Db 140 GluLysLysGlyLeuGluCysSerThrLeuTyrrArgThrGlnSerSerSerAsnProAla 159
 QY 565 GACTTAATGAGTGCCTTGAGAGCAGAAAGTCTCT----- 600
 Db :::: 111 :::: ::::
 Db 160 GluLeuArgGlnLeuLeuAspCysAspThrAlaSerLeuAspLeuGluMetPheAspVal 179
 QY 601 CATACACGAGAGAACCTGAATCAACTCATCTGTTAGAAAACCAAAATACTCAAGAGATT 660
 Db :::: 111 :::: 111111111 :::: 111
 Db 180 HisValLeuAlaAspAlaPheLysArgTyrLeuLeuAspLeuProAsn---ProValIle 198
 QY 661 CCACCTGCATTAACAGCTCT----- 681
 Db :::: 111111111 :::: ::::
 Db 199 ProValAlaValSerSerGlnLeuLeuIleSerLeuAlaProGluValGlnSerSerGlnGlu 218

CC STRAIN-BALB/C;
 RA MEDLINE-91191564; PubMed-1849460;
 RA Escobedo J.A.; Navankasatunas S.; Kavanaugh W.M.; Milfay D.,
 RA Fried V.A.; Williams L.T.;
 RT "cDNA cloning of a novel 85 kd protein that has SH2 domains and
 RT regulates binding of p13-kinase to the PDGF beta-receptor.";
 RL Cell 65:75-82(1991).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL: M60651; AAA39886.1; -
 DR HSSP: P23727; 1BFI.
 DR MGD: MG1:97583; PI3K1.
 DR InterPro: IPR001720; PI3kinase_P85.
 DR InterPro: IPR000198; RhogAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PR00678; PI3KINSEP85.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 3 79 SH3.
 FT DOMAIN 333 428 SH2 1.
 FT DOMAIN 624 718 SH2 2.
 SQ SEQUENCE 724 AA; 83414 MW; DEAA6EE297CF07A CRC64;

Alignment Scores:
 Pred. No.: 0 000845 Length: 724
 Score: 149.00 Matches: 76
 Percent Similarity: 32.76% Conservative: 39
 Best Local Similarity: 21.65% Mismatches: 96
 Query Match: 7.38% Indels: 140
 DB: 1 Gaps: 17

US-09-856-061-3 (1-1129) x p85A_MOUSE (1-724)

QY 421 CCTTACACACCTCCGCTCTCATACACTCCAGAGAGTACCAACCTTGCCCTT 480
 DB 85 ProthProlysrProargPro-----Proarg-----ProleupProval 97
 QY 481 GACCCGAGACGACGAG-----CCACCT 504
 DB 98 AlapProGlySerSerlysrThrglnAlaAspThrglnGlnAlaLeupProleupProasp 117
 QY 505 TTTATCTCAGAGACACACCTTCCAGAGATCCAGGATGCCAGATGACATTAAGG 564
 DB 118 LeuAlaGlnGlnPheAlaProProAspVal---AlaProProLeuLeuIleIleLeu 136

QY 565 GACTTAAGTGAGCTC-----CTTGAAGCAGAAAAAGTTCCATACACGAGAAACCT 618
 DB 137 GlnAlaIleGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 156
 QY 619 GAATCAATCAATCTGTGAGA-----AACCAATATCTCAAGAG 657
 DB 157 AsnProAlaGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 176
 QY 658 ATTCACCTTGCCCTTATGACATCTTTCATTCACAGACAGCAACCAAGCTGCAAAACAGA 717
 DB 177 IleAspValHisValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 195
 QY 718 GATCATAGAGAGGACGACGACCCCTGTCTCT-----750
 DB 196 -----ProValIleProValAlaValIleuIleuIleuIleuIleuIleuIleu 208
 QY 751 -----CAGAGATGACGACCTCCAGCAGCTGC-----777
 DB 209 LeuAlaGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 228
 QY 778 -----AGCCCTCAGCAAAATATACCTCCCTATAAATAC-----810
 DB 229 LeuProAsnIleProIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 248
 QY 810 -----810
 DB 249 LysLeuSerGlnAlaSerSerLysAsnLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 268
 QY 810 -----810
 DB 269 SerProValLeuPheArgPheProAlaAlaSerSerAspAsnThrglnHisLeuIleuIleu 288
 QY 811 -----ACAAGCTG-----819
 DB 289 AlaIleGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 308
 QY 820 ---AGACACACCTTCCCTC-----AAAAGCTCTGTATCAAAAGATGTC 858
 DB 309 ProLysProProLysProThrThrValAlaAsnAsnSerMetAsnAsnMetSerLeu 328
 QY 859 CAGCACAATGATGTCATGATGCAATACAGCCCGCAGCAGCTGCAAGGCAATTCATG 918
 DB 329 GlnAspAlaGlnIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 347
 QY 919 AAGGAGAACAGATGATGATGCTGTCGCGATGTCGCAAAATGCAAGAGAAAG 978
 DB 348 ArgAspThrAlaAspGlyThrPheLeuValArgAspAlaSerThrIleuIleuIleuIleuIleu 367
 QY 979 CCTATGTTTGGCTGTGTTTATGAGAACCAAGCTCATAGTAA---ATCCGCTTC 1035
 DB 368 ---TyrThrIleuIleu-----ArgLysGlyIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 383
 QY 1036 CTGAGAGCATACAGACTTCCCTGGGAGACAGACTCAGAGAGATGAGAACTTTGAT 1095
 DB 384 PheHisArgAspLysIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 1128
 QY 1096 TCAGTAGAGACATCAATCAACACATCAAGAT 1128
 DB 400 SerValValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 410

RESULT 6
 ID P85A.RAT STANDARD: PRT: 724 AA.
 AC 063787; 063790; P70544; 055085;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
 DE p85-alpha subunit) (PtdIns-3-kinase p85-alpha) (PI3K).
 GN RAKR1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

[illegible]

QY 694 AGCAACACAGTGTGCAAAACAGATCATAGA----- 726
 Db 276 ProlalalasereserapantthgluhlsleulellysalaValgluleulelleser 295
 QY 727 -----GGAGCATGCAGCCCTGTCCTCCAGAGATCCAGCCATCCAGCAGCTGC 777
 Db 296 AlaglturpsergluarglnProlalProlaleuProPolysProPolysProThr 315
 QY 778 AGCCCTCAGCAAAATATACGCGCCCTATTAATACAGAGCTGGACACACCTTTCCCAA 837
 Db 316 SerllealaasnserMer----- 322
 QY 838 AGGTCTGATGAAAGATGCTCCAGACATGATGATGAGATATCAGCCGCGCAG 897
 Db 323 ---AsnAsnsmetserleuglnAspAlaglturptryrpglyAspIleserArgly 341
 QY 898 GCACTGAGAGAGCATTATGAGAGAGACAGAGATGATGCTGCTCCGAGATTGT 957
 Db 342 GluValasnGlulys---LeuArgAspThrAlaAspGlyThrPheLeuValArgAspAla 360
 QY 958 TCACCAAAATCCAGAGAGAGCCCTATGCTTGTGCTGTTTATGAGAACAAAGTCTAC 1017
 Db 361 SerThrlysmethIsglyAsp---TyrThrLeuThrleu-----ArglysglyGly 376
 QY 1018 AATGTAAAT---ATCCGCTTCCTGAGAGATCAGCAGATTGCGCTGGGAGCAGACTC 1074
 Db 377 AsnAsnlyslleulellysllePhehisArgAspGlylyslrglyPheSerAspProleu 396
 QY 1075 AGAGGAGATGAGAGATTGATTGATGAGAGACATCATCGAACACTACAGAAAT 1128
 Db 397 -----ThrPheAsnSerValValGluleulelAsnHisTyrArgAsn 410

RESULT 7
 BTK_MOUSE STANDARD; PRT; 659 AA.
 ID BTK_MOUSE
 AC P35991; Q61365;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tyrosine-protein kinase BTK (EC 2.7.1.112) (Bruton's tyrosine kinase)
 DE (Agammaglobulinemia tyrosine kinase) (ATK) (B cell progenitor kinase) (BPK) (Kinase EMB)
 GN BTK OR BPK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 NC NCBITaxID=10090;
 ON [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93145329; Pubmed=8425221;
 RA Tsukada S., Saffran D.C., Rawlings D.J., Parolini O., Allen R.C., Klisak I., Sparkes R.S., Kubagawa H., Mohandas T., Quan S., Belmont J.W., Cooper M.D., Conley M.E., Witte O.N.;
 RT Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia.";
 RL Cell 72:279-290(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9326578; Pubmed=8476425;
 RA Yamada N., Kawakami Y., Kimura H., Fukumachi H., Baier G., Altman A., Kato T., Inagaki Y., Kawakami T.;
 RT "Structure and expression of novel protein-tyrosine kinases, Emk and Bmt, in hematopoietic cells";
 RL Biochem. Biophys. Res. Commun. 192:231-240(1993).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9508160; Pubmed=7989760;
 RA Sideras P., Mueller S., Shiels H., Jin H., Khan W.N., Nilsson L., Parkinson E., Thomas J.D., Branden L., Larsson L., Paul W.E., Rosen F.S., Alt F.W., Vetric D., Smith C.I.E., Xanthopoulos K.G.;
 RT "Genomic organization of mouse and human Bruton's agammaglobulinemia tyrosine kinase (Btk) loci.";

RL J. Immunol. 153:5607-5617(1994).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C129;
 RX MEDLINE=95352959; Pubmed=7626884;
 RA Oeltjen J.C., Liu X., Lu J., Allen R.C., Muzny D.M., Belmont J.W., Gibbs R.A.;
 RT "Sixty-nine kilobases of contiguous human genomic sequence containing the alpha-galactosidase A and Bruton's tyrosine kinase loci.";
 RL Mamm. Genome 6:334-338(1995).
 RN [5]
 RP VARIANT XID CYS-28.
 RX MEDLINE=93324903; Pubmed=8332901;
 RA Rawlings D.J., Saffran D.C., Tsukada S., Largepanda D.A., Grimaldi J.C., Cohen L., Mohr R.N., Bazan J.F., Howard M., Copeland N.G., Jenkins N.A., Witte O.N.;
 RT "Mutation of unique region of Bruton's tyrosine kinase in immunodeficient Xid mice.";
 RL Science 261:358-361(1993).
 CC -1- FUNCTION: PLAYS A CRUCIAL ROLE IN B-CELL ONTOGENY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC -1- PTM: Autophosphorylated on Tyr-223 and Tyr-551. The tyrosine phosphorylation of Tyr-223 may create a docking site for a SH2 containing protein.
 CC -1- DISEASE: DEFECTS IN BTK ARE THE CAUSE OF MURINE X-LINKED IMMUNODEFICIENCY (XID).
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: L08967; AAA37316.1; -
 DR EMBL: L10627; -; NOT_ANNOTATED_CDS.
 DR EMBL: L29788; AAA66943.1; -
 DR EMBL: U58105; AAB47246.1; -
 DR HSSP: Q06187; IBMN.
 DR MGD: MGI:88216; Btk.
 DR InterPro: IPR001562; BTK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00779; BTK; 1.
 DR PRINTS: PRO0452; SH3DOMAIN.
 DR PRINTS: PRO0402; TECBTDOMAIN.
 DR PRINTS: PRO0109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00107; BTK; 1.
 DR SMART: SM00233; PH; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00326; SH3; 1.
 DR SMART: SM00219; TYKCa; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.
 DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.


```

Query Match: 6.93% Indels: 168
DB: 1 Gaps: 18
US-09-856-061-3 (1-1129) x 3BP2_MOUSE (1-559)

QY 166 CTTGGATGGAAGACATGGCAGTAAATTTTACGACCCGCGCTATAAGAA 225
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 82 MetArgAlaLagLugLutThrThrSerAsnValPheProPheLysIleIleHis 101
QY 226 TCGAATATGAGATACACATTTTCAGAGTTGCAATGAC----- 267
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 102 SerLysLysHisArgThrThrPhePheSerAlaSerSerGluAspGluArgLysSer 121
QY 268 -----ACGCCCTCCGCTAGAC 285
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 122 MetAlaPheValArgArgLugLutIleGlyHisPheHisGluLysLysGluLeuProLeuAsp 141
QY 286 ACCAGAGCTTATGTCATTTGACAGCCGACCTGGACACACAGACAGAGTTGAAAGA 345
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 142 ThrSerAspSerSer-----AspThrAspSerPheTyrgLysAla 155
QY 346 GTGGACAACCCATTCCAGGAGCTCAGACGCCAAACATTAAAGAGATGATCCGTA 405
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 156 ValGluArgProIleAspIleSerLeuSerSerTyrrProMetAspAsnGluAspTyrgL 175
QY 406 AGAAGACAGATTCCT-----TTACCACTCTCGGCT----- 441
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 176 HisGluAspGluAspSerTyrrLeuGluProAspSerProGlyProMetLysLeuGlu 195
QY 442 ---CTATACACTTCGAGAAAGTACCAACCTTCCCTCGGCGGAGCGAGACAGCAG 498
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 196 AspAlaLeuThrTyrrProAlaTyrrProPro---ProValProValProAlaGlyLys 214
QY 499 CCACCTTATCT-----CAGACACACACTT----- 525
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 215 ProAlaPheSerAspLeuProAlaGluHisSerPheThrSerLysSerProSerProLeu 234
QY 526 -----CCAGACTCCAGGAATGCCAGTCAAGTAAAGCTTA---AGGACTTA 570
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 235 LeuProProProProProLysArgGlyLeuProAspThrGlySerAlaProGluAspAla 254
QY 571 AGTGAAGTCTTGAACACAGAAAGTCTCATACACAGAGAGGCTGATACACTCAT 630
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 255 LysAspAlaLeuGlyLeuArgArgValGluProGlyLeuArgValProAlaThrProArg 274
QY 631 CTGTTGAACAAACCAATACTCAAGATTCCTGCACTGCACTTATGAC----- 675
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 275 ArgMetSerAspProProMetSerAsnValProThrValProAsnLeuArgLysHisPro 294
QY 675 ----- 675
DB 295 CysPheArgAspSerValAsnProGlyLeuGluProThrProGlyHisGlyThrSer 314
QY 676 -----AGTTCTCATTCACAGACACACACAGTGTGCAAAACAGAGAT----- 720
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 315 SerValSerSerSerThrThrMetAlaValAlaThrSerArgAsnGlySasplysLeuLys 334
QY 721 -----CATAGAGAGGAGCATGACCC----- 741
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 335 SerPheHisLeuSerSerArgLysProProThrSerGluProProProValProAlaAsn 354
QY 742 -----TGTTCTCTCAGAGATGCCAG----- 762
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 355 LysProLysPheLeuLysIleAlaGluLupProSerProArgLysAlaAlaLysPheAla 374
QY 763 -----CCTCCAGCCAGCTGCACAGCCATGCAAAATATCTGCTATATAATACACAGC 816
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 375 ProValProProValAlaProArgProProValGluLysMetProMetProGluAlaThr 394
QY 817 TGGAGACACACTTCCCAAAAGTGTGATAGAAAGATGTCACAGAC----- 864
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 395 ValArgProAlaValLeuProArgProGluAsnThrProLeuProHisLeuGluArgSer 414

```

```

QY 864 ----- 864
DB 415 ProProAspGlyLysInserPheArgGlyPheSerPheGluLysAlaArgInProSerGln 434
QY 865 -----AATGAA 870
DB 435 AlaAspThrGlyLysLysAspSerAspGluAspTyrgLysValProLeuProAsnSer 454
QY 871 TGGTACATTGGAGAAATACAGCCCGCAGCGAGTGGAAGAGCATTCATGAAAG----- 921
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 455 ValPheValAsnThrThrGluSerCysGluValArgLeuPheLysAlaThrAspPro 474
QY 922 -----GAGAACAGAGATGCTACTTCTGTCGAGATGTTGCCAAATCCAGAGAA 975
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 475 ArgGlyGluProGluAspGlyLeuTyrrCysIleArgAsnSerThrLysSerGlyLys 494
QY 976 GAGCCCTATGTTTGGCTGTGTTTATGAG-----AACAACTCTACATGTAATAATC 1029
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 495 -----ValLeuValValTrpAspGluSerSerAsnLysValArgAsnTyrrArgIle 511
QY 1030 CGCTTCGTGAGAGGAAATCAGCAGTTGCCCTGGAGACAGACTCAGAGAGATGGAAG 1089
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 512 -----PheGluLysAspSerLysPheTyrr-----LeuGluGlyGluValLeu 525
QY 1090 TTGTATTCAGTAGACATCATCATGAACTAC 1122
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
DB 526 PheAlaSerValGlySerMetValGluHisTyrr 536

RESULT 9
BTK_HUMAN
ID BTK_HUMAN STANDARD; PRT; 659 AA.
AC 006187;
DT 01-JUN-1994 (Rel. 29, last sequence update)
DT 01-JUN-1994 (Rel. 41, last annotation update)
DE 15-JUN-2002 (Rel. 41, last annotation update)
DE Tyrosine-protein kinase BTK (EC 2.7.1.112) (Bruton's tyrosine kinase)
DE (Agammaglobulinemia tyrosine kinase) (ATK) (B cell progenitor
DE kinase) (BPK).
GN BTK OR ATK OR AGMX1 OR BPK.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93140868; PubMed=8380905;
RA Vetric D., Vorechovsky I., Sideras P., Holland J., Davies A.,
RA Flinter F., Hammarstrom L., Kinnon C., Levinsky R.J., Bobrow M.,
RA Smith C.I.E., Bentley D.R.;
RT "The gene involved in X-linked agammaglobulinemia is a member of the
RT src family of protein-tyrosine kinases.";
RL Nature 361:226-233(1993).
RN [2]
RP ERRATUM.
RA Vetric D., Vorechovsky I., Sideras P., Holland J., Davies A.,
RA Flinter F., Hammarstrom L., Kinnon C., Levinsky R.J., Bobrow M.,
RA Smith C.I.E., Bentley D.R.;
RL Nature 364:362-362(1993).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=94377492; PubMed=8090769;
RA Ohta Y., Haire R.N., Litman R.T., Fu S.M., Nelson R.P., Kratz J.,
RA Kornfeld S.J., la Morena M., Good R.A., Litman G.W.;
RT "Genomic organization and structure of Bruton agammaglobulinemia
RT tyrosine kinase: localization of mutations associated with varied
RT clinical presentations and course in X chromosome-linked
RT agammaglobulinemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:9062-9066(1994).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=95012452; PubMed=7927535;
RA Rohrer J., Parolini O., Belmont J.W., Conley M.E.;

```

- RT "The genomic structure of human BTK, the defective gene in X-linked
RT agammaglobulinemia.";
RL Immunogenetics 40:319-324(1994).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE-95352959; PubMed-7626884;
RA Oeltjen J.C., Liu X., Lu J., Allen R.C., Muzny D.M., Belmont J.W.,
RA Gidbs R.A.;
RT "Sixty-nine kilobases of contiguous human genomic sequence containing
RT the alpha-galactosidase A and Bruton's tyrosine kinase loci.";
RL Mamm. Genome 6:334-338(1995).
RN [6]
RP SEQUENCE FROM N.A.
RA Oeltjen J.C., Malley T.M., Muzny D.M., Miller W., Gidbs R.A.,
RA Belmont J.W.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE OF 1-442 FROM N.A.
RX MEDLINE-93145329; PubMed-8425221;
RA Tsukada S., Saffran D.C., Rawlings D.J., Parolini O., Allen R.C.,
RA Kilsak I., Sparks R.S., Kubagawa H., Mohandas T., Quan S.,
RA Belmont J.W., Cooper M.D., Conley M.E., Witte O.N.;
RT "Deficient expression of a B cell cytoplasmic tyrosine kinase in
RT human X-linked agammaglobulinemia.";
RL Cell 72:279-290(1993).
RN [8]
RP PHOSPHORYLATION OF GTF2I, AND MUTAGENESIS.
RX MEDLINE-97165069; PubMed-9012831;
RA Yang W., Desiderio S.;
RT "BAP-135, a target for Bruton's tyrosine kinase in response to B cell
RT receptor engagement.";
RL Proc. Natl. Acad. Sci. U.S.A. 94:604-609(1997).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 1-170.
RX MEDLINE-97361821; PubMed-9218782;
RA Hyvoenen M., Saraste M.;
RT "Structure of the PH domain and Btk motif from Bruton's tyrosine
RT kinase: molecular explanations for X-linked agammaglobulinemia.";
RL EMBO J. 16:3396-3404(1997).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-170.
RX MEDLINE-99216540; PubMed-10196129;
RA Baraldi E., Carugo K.D., Hyvoenen M., Surdo P.L., Riley A.M.,
RA Potter B.V.L., O'Brien R., Ladbury J.E., Saraste M.;
RT "Structure of the PH domain from Bruton's tyrosine kinase in complex
RT with inositol 1,3,4,5-tetrakisphosphate.";
RL Structure 7:449-460(1999).
RN [11]
RP STRUCTURE BY NMR OF 209-275.
RX MEDLINE-98153176; PubMed-9485443;
RA Hansson H., Mattsson P.T., Allard P., Haapaniemi P., Vihinen M.,
RA Smith C.I.E., Heerd T.;
RT "Solution structure of the SH3 domain from Bruton's tyrosine kinase.";
RL Biochemistry 37:2912-2924(1998).
RN [12]
RP REVIEW ON XLA VARIANTS.
RX MEDLINE-96174621; PubMed-8594569;
RA Vihinen M., Iwata T., Kinnon C., Kwan S.-P., Ochs H.D.,
RA Vorechovsky I., Smith C.I.E.;
RT "Forkbase, mutation database for X-linked agammaglobulinemia (XLA).";
RL Nucleic Acids Res. 24:160-165(1996).
RN [13]
RP REVIEW ON XLA VARIANTS.
RX MEDLINE-97169387; PubMed-9016530;
RA Vihinen M., Belohradsky B.H., Halre R.N., Hollnaki-Peder E.,
RA Kwan S.-P., Lappalainen I., Lehtvaesalho H., Lester T., Melndl A.,
RA Ochs H.D., Ollila J., Vorechovsky I., Weiss M., Smith C.I.E.;
RT "Forkbase, mutation database for X-linked agammaglobulinemia (XLA).";
RL Nucleic Acids Res. 25:166-171(1997).
RN [14]
RP VARIANTS XLA TRP-288; GLY-307; ASP-607 AND S-V-F-S-S-T-R-103 INS.
RX MEDLINE-94214481; PubMed-8162056;
RA Bradley L.A.D., Sweatman A.K., Lovering R.C., Jones A.M., Morgan G.,
RA Levinsky R.J., Kinnon C.;
RT "Mutation detection in the X-linked agammaglobulinemia gene, BTK,
RT using single strand conformation polymorphism analysis.";
RL Hum. Mol. Genet. 3:79-83(1994).
RN [15]
RP VARIANTS XLA HIS-28 AND TRP-288.
RX MEDLINE-94214435; PubMed-8162018;
RA de Weers M., Mensink R.G.J., Kraakman M.E.M., Schuurman R.K.B.,
RA Hendriks R.W.;
RT "Mutation analysis of the Bruton's tyrosine kinase gene in X-linked
RT agammaglobulinemia: identification of a mutation which affects the
RT same codon as is altered in immunodeficient xid mice.";
RL Hum. Mol. Genet. 3:161-166(1994).
RN [16]
RP SEQUENCE FROM N.A., AND VARIANTS XLA S-334, R-506, Q-520, W-562 AND
RP K-630.
RX MEDLINE-95152493; PubMed-7880320;
RA Hagemann T.L., Chen Y., Rosen F.S., Kwan S.-P.;
RT "Genomic organization of the Btk gene and exon scanning for mutations
RT in patients with X-linked agammaglobulinemia.";
RL Hum. Mol. Genet. 3:1743-1749(1994).
RN [17]
RP VARIANTS XLA D-113; C-361; Q-520; P-542; W-562; K-630 AND P-652.
RX MEDLINE-95152494; PubMed-7849697;
RA Conley M.E., Fitch-Hilgenberg M.E., Cleveland J.L., Parolini O.,
RA Rohrer J.;
RT "Screening of genomic DNA to identify mutations in the gene for
RT Bruton's tyrosine kinase.";
RL Hum. Mol. Genet. 3:1751-1756(1994).
RN [18]
RP VARIANTS XLA H-28; P-33; P-408; G-589; D-613 AND 260-Q--E-280 DEL.
RX MEDLINE-95152522; PubMed-7849721;
RA Zhu Q., Zhang M., Winkelsstein J., Chen S.-H., Ochs H.D.;
RT "Unique mutations of Bruton's tyrosine kinase in fourteen unrelated
RT X-linked agammaglobulinemia families.";
RL Hum. Mol. Genet. 3:1899-1900(1994).
RN [19]
RP VARIANTS XLA E-430; Q-520; Q-525; P-562; V-582; G-589; E-594 AND
RP D-613.
RX MEDLINE-95108046; PubMed-7809124;
RA Vihinen M., Vetric D., Maniar H.S., Ochs H.D., Zhu Q., Vorechovsky I.,
RA Webster A.D.B., Notarangelo L.D., Nilsson L., Sowadski J.M.,
RA Smith C.I.E.;
RT "Structural basis for chromosome X-linked agammaglobulinemia: a
RT tyrosine kinase disease.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:12803-12807(1994).
RN [20]
RP VARIANT XLA PHE-64, AND CHARACTERIZATION OF OTHER XLA VARIANTS.
RX MEDLINE-95151728; PubMed-7849006;
RA Vihinen M., Zvelebil J.J.M., Zhu Q., Broolmans R.A., Ochs H.D.,
RA Zegers B.J.M., Nilsson L., Waterfield M.D., Smith C.I.E.;
RT "Structural basis for pleckstrin homology domain mutations in
RT X-linked agammaglobulinemia.";
RL Biochemistry 34:1475-1481(1995).
RN [21]
RP VARIANTS XLA S-25; W-288; M-370; V-509; P-525; K-526; V-582
RP AND R-594.
RX MEDLINE-95227177; PubMed-7711734;
RA Vorechovsky I., Vihinen M., de Saint Basile G., Honsova S.,
RA Hammarstrom L., Mueller S., Nilsson L., Fischer A., Smith C.I.E.;
RT "DNA-based mutation analysis of Bruton's tyrosine kinase gene in
RT patients with X-linked agammaglobulinemia.";
RL Hum. Mol. Genet. 4:51-58(1995).
RN [22]
RP VARIANTS XLA IYS-567, LEU-587 AND HIS-641.
RX MEDLINE-95359977; PubMed-7633420;
RA Jin H., Webster A.D.B., Vihinen M., Sideras P., Vorechovsky I.,
RA Hammarstrom L., Bernatowska-Matuszkiewicz E., Smith C.I.E.,
RA Bobrow M., Vetric D.;
RT "Identification of Btk mutations in 20 unrelated patients with
RT X-linked agammaglobulinemia (XLA).";
RL Hum. Mol. Genet. 4:693-700(1995).
RN [23]

Best Local Similarity: 20.55% Mismatches: 172
 Query Match: 6.73% Indels: 142
 DB: 1 Gaps: 20

US-09-856-061-3 (1-1129) x 3BP2_HUMAN (1-561)

```

OY 22 AAAAAAGGTCATGCTGCTGCAATAGTCCACAGGCCATACCAGAGATGAACAAG 81
DB 104 LysHisArgThrTrp-----PhePheSerAlaSerSerGluGluArg----- 118
OY 82 CCTCTTCAAGCTGGGAAAGAACTTGTGTCAGCTCCCTGATGAGCAAAAGCCACACT 141
DB 119 -----LysSerTrpMetAlaLeuLeuArgGluIleGluIlePhe 132
OY 142 GATGATGATATGATGACCTGAGCTGCGATGAGAGAGACATGGCAGTGAATTAAT 201
DB 133 HisGluIleLysArgLeuProLeuAspThrSerAspSerSerSerAspThrSerPhe 152
OY 202 TTACAGCC-----CGGCTATA-----ANG 222
DB 153 TyrGluAlaValAlaGluArgProValAspIleSerLeuSerProTyrProThrAspAsnGlu 172
OY 223 GAATGTGATATGATGATACACATATTTCAGGTTCAGATGACACATCCCTCCGTTA 282
DB 173 AspTyrGluHisAspAspSerLysAspSerTyrLeuGluProAspSerProLysPro 191
OY 283 GACACAGAGACCTTATCTCCATGACAGCCAGCTGGACACACAGAGAGTTGGA 342
DB 192 ---GluTyrGluGluAspAlaLeuMetHisProProAlaTyrProProProValPro 210
OY 343 AGAGTGGACAAACCCATTTCACAGGAGCTC-----AGAACCCAAACATTAAAGA 393
DB 211 ThrProArgLysArgProAlaPheSerAspMetProArgAlaHisSerPheThrSerGly 230
OY 394 GATGATCGGTGAAGAAAGAAAGATTCCTTACCAACCTCCGCG----- 438
DB 231 ProGlyProLeu-----LeuProProProProProIleGlyLeuProAsp 246
OY 439 -----CCTCTCATTAACCTTCGGAAGATAC 465
DB 247 ValGlyLeuAlaIleAlaGluAspSerLysAspAspProLeu-----CysProArgArgAla 264
OY 466 CAACCTTGCCCTGAGCCGCGAGAGACAG-----CCACCTTATCTACG 513
DB 265 GluProCysProArgValAlaThrProArgArgMetSerAspProLeuSer 283
OY 514 AGACACACTTCCAGAGTCCAGGGAAG----- 543
DB 284 -----ThrMetProThrAlaProGlyLeuArgLysProProCysPheArgGluSerAla 301
OY 544 -----CCAGTCCAGATTAAGCTTAAGGAGCTTAAGTAGAGTGC 579
DB 302 SerProSerProGluProTrpThrProGlyHisGlyAlaCysSerThrSerSerAlaIle 321
OY 580 CTGTAAGCAAGAAAGTTCTCATACCAAGAGAAAGCTGAATCAATCATCTGTATTAGA 639
DB 322 IleMetAlaThrAlaThrSerArgAsnGlySerLysLeuLysSerPheHisLeuSerPro 341
OY 640 AAC-----CAAAATACCTCAAGATTCACCTTGCATTAAGAGTTCTTATTCACGACA 693
DB 342 ArgGlyProProTrpThrSerGluProProProValProAlaAsnLysProLysPhe 359
OY 694 AGCAACACAGTGTGCAAAACAGAGATCATAGAGGAGCATCAGCCCTGTTCTCTCAG 753
DB 360 -----LeuLysIleAlaGluLysProProProArgGluAlaIleMetProGly 375
OY 754 AGATGCCAGCTTCAGCCAGCTGACGCCCTCAGAAATATATGCTTAAATACACA 813
DB 376 LeuPheValProProValAlaProArgProProAlaLeuLysLeuProValProGluAla 395
OY 814 AGCTGAGACCACTTCCCAAAAGGCTGTATAGAAAGATGTCACAGAC----- 864
DB 396 MetaLysProAlaValLeuProArgProGluLysProGluLysProGluHisLeuGlnArg 415

```

```

OY 864 ----- 864
DB 416 SerProProAspGlyGlnSerPheArgSerPheSerPheGluLysProArgGlnProSer 435
OY 865 -----AAT 867
DB 436 GlnAlaAspThrGlyGlyAspAspSerAspGluAspTyrGluLysValProLeuProAsn 455
OY 868 GATGGTACATGATGGAATATACGCCCGCAGGAGGAAAGCAAGCATTCATGAA----- 921
DB 456 SerValPheValAsnThrThrGluSerCysGluValGluArgLeuPheLysAlaThrSer 475
OY 922 -----GAGAACAGGATGATGATTTCTTGTCGCGAGATTTCCCAAAATCCAAAG 972
DB 476 ProArgGlyGluProGlnAspGlyLeuTyrCysIleArgAsnSerSerThrLysSerGly 495
OY 973 GAAGAGCCCTATGTTTGGCTGTGTTTATGAG-----ACAAGTCTACATGTAA 1026
DB 496 Lys-----ValLeuValAlaTrpAspGluThrSerAsnLysValArgAsnTyrArg 512
OY 1027 ATCCGCTTCCTGAGAGGAATACAGAGTTGCCCTGGGAGACGACTCAGAGAGATGAG 1086
DB 513 Ile-----PheGluLysAspSerLysPheTyr-----LeuGluGlyGluVal 526
OY 1087 AAGTTGATTGATCAGTAGAAGACATCATCGAACACTAC 1122
DB 527 LeuPheValSerValGlySerMetValGluHisTyr 538

```

RESULT 12

P85B_BOVIN STANDARD: PRT; 724 AA..

ID P23726;

AC 01-NOV-1991 (Rel. 20, Last sequence update)

DT 01-NOV-1991 (Rel. 20, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase

DE p85-beta subunit) (PtdIns-3-kinase p85-beta).

GN PIK3R2.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=91191567; PubMed=1707345;

RA Otsu M., Hiles I.D., Goot I., Fry M.J., Ruiz-Larrea F., Panayotou G.,

RA Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,

RA Courtenidge S.A., Parker P.J., Waterfield M.D.;

RT "Characterization of two 85 kd proteins that associate with receptor

RT tyrosine kinases, middle-T/pp60c-src complexes, and p13-kinase.";

RL Cell 65:91-104(1991).

RN [2]

RP CIRCULAR DICHOISM AND FLUORESCENCE SPECTROSCOPY.

RX MEDLINE=93049176; PubMed=1330535;

RA Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski B., Dhand R.,

RA Fry M.J., Blundell T.L., Wolmer A., Waterfield M.D.;

RT "Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2

RT domain with a PDGF receptor phosphorylation site: structural features

RT and analysis of conformational changes.";

RL EMBO J. 11:4261-4272(1992).

CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE

CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING

CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.

CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)

CC SUBUNITS.

CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.

CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.

CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.

CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

```

QY      301  TCCATTGGA---CAGCCGACCTGGACAACACACACGAGGTTGGAAAGAGTGGCAAAACC 357
Db      53  SerValGlyTrpMetProGlyLeuAsnGluArgHsrGly---GlnArgGlyAspPhePro 71
QY      358  ATTTCACAGGACGTCACAGACCAAAACATTAAAGAGACATGCATCCGTAAAGAAACACAG 417
Db      72  GlyThr-----TyrValGluPheLeuGlyProValAlaIleuAlaIArg----- 85
QY      418  ATTCCATTACCACTCGTCGGCCCTCATTAACACTTCGGAAGAAGTACCAACCTTGCCC 477
Db      86  ---ProGlyProAlaGlyProAlaGlyGlyProArgProLeuProAlaArgProAlaGlyPro 104
QY      478  CCTGACCGGAGACGACGAGGCCACTTTATCTTCAGACACACCTTTCCAGAAGTCAG 537
Db      105  ProGluProGlyLeuThrLeuProAspLeuProGluInPheSerProProAspVal--- 123
QY      538  GGATGCCGACGTACGATTAAGCTTAAGGACTTAAGTAGCTCCTTGAGACAGAAAAGTT 597
Db      124  AlaProPheIleuValLysLeu-----ValGluAlaIleGluArgThr 138
QY      598  CCTCATACACGAGAGAGCCGATCAACTCATCTGTAAACAAACAAATATCTCAAG 657
Db      139  GlyLeuAspSerTyrArgProGluInProProAlaValAlaArgHsrAspTrpSerLeuSerAsp 158
QY      658  ATTCCA-----CTTGCCATT 672
Db      159  ValGluInTrpAspAlaAlaAlaLeuSerAspGlyValLysGlyPheLeuLeuAlaLeu 178
QY      673  AGCAGTCTTCATTACGACAGACAGAACACACAGTGTGCAAAACAGAGATCATATAGGAGGC 732
Db      179  ProAlaProLeuValThrProGluAlaIleAlaIleValAlaHisAspAlaLeuArgGluAla 198

```

QY	733	ATGGAGCCCGTTCCTCCAG-----	753
Db	199	AlaGlyProValGlyProAlaLeuGluProProThiLeuProLeuHisHisAlaLeuThr	218
QY	753	-----	753
Db	219	LeuArgPheLeuLeuGlnHisLeuGlyArgValAlaGlyAlaArgAlaProAlaProGlyPro	238
QY	754	-----AGATGCCAGCCTCCAGCC	771
Db	239	AlaValArgAlaLeuGlnGlyAlaThrPheGlyProLeuLeuLeuArgAlaProProProPro	258
QY	772	AGCGAGCGCCCT-----	783
Db	259	SerProProProGlyGlyAlaProAspGlyThrGluProThrProAspPheProAlaLeu	278
QY	784	-----CAGAAATGTACTGCCCTATAAATAC	810
Db	279	LeuValGluIysLeuLeuGlnGlnHisLeuGluGluGlnGlnValAlaProProAlaLeu	298
QY	811	ACAAGCTGAGACCACCTTTCCCAAAAGTCT-----	843
Db	299	-----ProPolsProPolsThyLysProAlaProThrGlyLeuAlaAsn	314
QY	844	-----GATGAAGAAGATGTCGACGACCAATGATGGTATGCTGGAGATAATGACGCGCCAG	897
Db	315	GlyGlySerProProSerLeuGlnAspAlaGluThrPryThrPrelYaspIseArgGlu	334
QY	898	GCACTGAGAGAGCCATCATGAGAAGAGAACAGAGTGATGTTCTTGTCGAGATGT	957
Db	335	GluValAsnGluLys---LeuArgAspThrProAspGlyThrPheLeuValAlaArgAspAla	353
QY	958	TCCACAAATCCAAAGAGAGCCCATGTTTGCGCTGTGTTTATGAAACAAGTCTAC	1011
Db	354	SerSerLysIleGlnGlyGlu---TyrThrLeuThrLeu-----ArgLysGlyGly	369
QY	1018	AATGTAAA---ATCCGCTTCCTCGAGAGAAATCAGAGATGTTGCCCTGGGACAGGACATC	1074
Db	370	AsnAsnLysLeuIleLysValAlaPheHisArgAspGlyHisIleGlyThrPheSerGluProLeu	389
QY	1075	AGAGAGATGAGAGATTGATTCAGTAGAAGACATCATGCAACCTACAGAAT	1128
Db	390	-----ThrPheCysSerValValAspLeuIleThrHisIleArgHis	403
RESULT 13			
P55G_HUMAN STANDARD; PRT; 461 AA.			
AC	092569; 060482;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	phosphatidylinositol 3-kinase regulatory gamma subunit (PI3-kinase		
DE	p85-gamma subunit) (Pcdins-3-kinase p85-gamma) (p55PIK).		
GN	PIK3R3.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RA	Suzuki T.;		
RT	"Molecular cloning of human p55piK."		
RL	Submitted (OCT-1996) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).		
RC	TISSUE=Fetal brain;		
RX	MEDLINE=98192526; PubMed=9524259;		
RA	Day B. R., Furlanetto R.W., Nissley S. P.;		
RT	"Cloning of human p55 gamma, a regulatory subunit of		
RT	phosphatidylinositol 3-kinase, by a yeast two-hybrid library screen		
RT	with the insulin-like growth factor-I receptor."		
RL	Gene 209:175-183(1998).		

CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES THROUGH ITS SH2 DOMAIN AND REGULATES THEIR KINASE
 CC ACTIVITY. DURING INSULIN STIMULATION, IT ALSO BINDS TO IRS-1.
 CC -1- SUBUNIT: HETEROIDIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN BRAIN AND TESTIS. LOWER
 CC LEVELS IN ADIPOSE TISSUE, KIDNEY, HEART, LUNG AND SKLETERAL MUSCLE.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: HIGH, WITH OTHER P85 SUBUNITS, AND WITH P85-ALPHA AND
 CC P85-BETA SUBUNITS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D88532; BA13636.1; -
 CC EMBL: AF028785; AAC39696.1; -
 CC HSSP: P23727; 2PMB.
 CC Genew: HGNC:8981; PIK3R3.
 CC MIM: 606076; -
 CC InterPro: IPR001720; PI3kinase_P85.
 CC InterPro: IPR000980; SH2.
 CC Pfam: PF00017; SH2; 2.
 CC PRINTS: PRO0678; PI3KINSEP85.
 CC PRINTS: PR00401; SH2DOMAIN.
 CC ProDom: PD000093; SH2; 2.
 CC SMART: SM00252; SH2; 2.
 CC PROSITE: PS50001; SH2; 2.
 CC SH2 domain; Repeat; Phosphorylation; Alternative splicing.
 CC KW DOMAIN 34 44
 CC FT DOMAIN 65 160 SH2 1.
 CC FT DOMAIN 358 452 SH2 2.
 CC FT MOD_RES 341 341 PHOSPHORYLATION (BY SIMILARITY).
 CC FT VARSPLIT 36 71 MISSING (IN ISOFORM 3).
 CC FT VARSPLIT 256 314 MISSING (IN ISOFORM 2 AND ISOFORM 3).
 CC FT CONFLICT 21 21 P -> L (IN REF. 2).
 CC SO SEQUENCE 461 AA; 54462 MW; C03ECFE22BEB5089 CRC64;
 CC -----
 CC Alignment Scores:
 CC Pred. No.: 0.0308 Length: 461
 CC Score: 128.00 Matches: 49
 CC Percent Similarity: 45.64% Conservative: 19
 CC Best Local Similarity: 32.89% Mismatches: 45
 CC Query Match: 6.34% Indels: 36
 CC DB: 1 Gaps: 10
 CC -----
 CC US-09-856-061-3 (1-1129) x P55G_HUMAN (1-461)
 CC QY 718 GATCATAGAGGAGGAGCCCTGTCTCTCTAG-----AGATGCCAG 762
 CC DB 14 AsPTpArGgLUvAlMeKMePrOtyrSerThrGluLeuIlePheTyrILeGLuMetasp 33
 CC QY 763 CCTCAGCAGCAGCTGCAGCCCTCAGCAAAATATACCTGATTAATATACCAAGCTGGAGA 822
 CC DB 34 PropProlA-----LeuPropProlys----- 40
 CC QY 823 CCACCTTCCCAAAAGGCT-----GATAGAAGGAT-----GTC 858
 CC DB 41 PropProlysPrometThSerAlaValProAsnglyMetLysAsPserSerValSerIeu 60
 CC QY 859 CAGCAAAATGATGTGATATGAGATATACAGCCCGCAGGAGCGAGGAGGATTCATG 918
 CC DB 61 GlAsPAlaGluTPrTyTIPeGLysPLeSerArGluGLuValAsnAsPlys---Leu 79
 CC QY 919 AAGGAGAAACAAGATGATGTTCTTGGTCGAGATGTGTTCCACAAATTCACAGGAGAG 978
 CC DB 80 ArGAsPMeLProAsPGLyThrPheLeuValArGAsPAlaSerThrLysMeLInGLyasp 99

QY 979 CCTATGTTGGCGTGTGTTTATGAGAACAAAGTCATCATGTAATA---ATCGGCTC 1035
 DB 100 ---TyrThrLeuThrIeu-----ArgLysGLyGLysAsnAsPLeuIleLysIle 115
 QY 1036 CTGGAGAGGAATCAGCAGTGTGCTGGGAGCAGCATCGAGAGATGAGAGATTGAT 1095
 DB 116 TyrHlAsrGAsPGLyLysTyrcILyPheSerArProLeu-----ThrPheAsn 131
 QY 1096 TCAGTAAAGACATCATCATCAACACTAC 1122
 DB 132 SerValValGLuLeuIleAsnHlStyr 140
 CC RESULT 14
 CC P85_MOUSE
 CC ID P85_MOUSE STANDARD; PRT; 722 AA.
 CC AC 008908;
 CC DT 15-DEC-1998 (Rel. 37, Created)
 CC DT 15-DEC-1998 (Rel. 37, Last sequence update)
 CC DT 15-JUN-2002 (Rel. 41, Last annotation update)
 CC DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 CC P85-beta subunit) (Ptdins-3-kinase p85-beta).
 CC GN PIK3R2.
 CC OS Mus musculus (Mouse).
 CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 CC OX NCBI_Taxid=10090;
 CC RN [1]
 CC RP SEQUENCE FROM N.A.
 CC RC STRAIN=NIH SWISS;
 CC RX MEDLINE=98241181; PubMed=9582025;
 CC RA Janssen J.W.G., Schleithoff L., Bartam C.R., Schulz A.S.;
 CC RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase
 CC RT p85beta subunit and HUNKOR8, a putative deubiquitinating enzyme.";
 CC RL Oncogene 16:1767-1772(1998).
 CC CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC -1- SUBUNIT: HETEROIDIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y13569; CAA73903.1; -
 CC HSSP: P23727; 2PMB.
 CC MGD: MGI:1098772; PIK3R2.
 CC InterPro: IPR001720; PI3kinase_P85.
 CC InterPro: IPR000198; RHO-GAP.
 CC InterPro: IPR000980; SH2.
 CC InterPro: IPR001452; SH3.
 CC Pfam: PF00017; SH2; 2.
 CC Pfam: PF00018; SH3; 1.
 CC Pfam: PF00620; RHO-GAP; 1.
 CC PRINTS: PRO0678; PI3KINSEP85.
 CC PRINTS: PR00401; SH2DOMAIN.
 CC ProDom: PD000066; SH3; 1.
 CC ProDom: PD000093; SH2; 2.
 CC SMART: SM00324; RHO-GAP; 1.
 CC SMART: SM00326; SH2; 2.
 CC SMART: SM00326; SH3; 1.
 CC PROSITE: PS50001; SH2; 2.
 CC PROSITE: PS50002; SH3; 1.
 CC SH3 domain; SH2 domain; Repeat.

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirsky R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yen R.-F., Zavel J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RL "The genome sequence of *Drosophila melanogaster*.";
CC Science 287:2185-2195(2000).
CC -1- FUNCTION: ADAPTER PROTEIN WHICH INTERACTS WITH C-TERMINAL PORTION
CC OF MRC, HOMOLOG OF HUMAN DOCK180. MAY PLAY A ROLE IN CELLULAR
CC PROCESSES THROUGHOUT DEVELOPMENT.
CC -1- TISSUE SPECIFICITY: EMBRYONIC ZYGOTIC EXPRESSION IS SEEN IN
CC INVAGINATING PRESUMPTIVE MESODERM AND ECTODERMALLY DERIVED TISSUES
CC DURING GASTRULATION. AT STAGE 8, EXPRESSION IS ALSO SEEN IN
CC ANTERIOR AND POSTERIOR MIDGUT AND CEPHALIC FURROW. BY STAGE 9,
CC EXPRESSION IS HIGHEST IN VISCERAL MESODERM OF ANTERIOR AND
CC POSTERIOR MIDGUT, VENTRAL NERVE CORD AND SOMATIC MESODERM.
CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY
CC THROUGHOUT EMBRYOGENESIS, DECLINES DURING LARVAL STAGES AND
CC REAPPEARS DURING PUPATION.
CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC -----
CC "This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE003844; AAF59362.1; -
DR EMBL: AF112976; AAD28428.1; -
DR HSSP: O64010; ICKA.
DR FlyBase: FBgn002481; Crk.
DR InterPro: IPR000980; SH2.
DR InterPro: IPR001452; SH3.
DR Pfam: PF00017; SH2; 1.
DR Pfam: PF00018; SH3; 1.
DR PRINTS: PRO0401; SH2DOMAIN.
DR PRINTS: PRO0452; SH3DOMAIN.
DR PRODOM: PD000066; SH3; 1.
DR PRODOM: PD000093; SH2; 1.
DR SMART: SM00252; SH2; 1.
DR SMART: SM00326; SH3; 2.
DR PROSITE: PS50001; SH2; 1.
DR PROSITE: PS50002; SH3; 1.
DR SH2 domain; SH3 domain; Repeat: Developmental protein.
KW SH2 DOMAIN 12 114 SH2.
FT DOMAIN 117 165 SH3 1.
FT DOMAIN 220 259 SH3 2.
SQ SEQUENCE 271 AA: 31205 MW: 41847.43 IS0932DC CRC64:

Alignment Scores:
Pred. No.: 0.0342 Length: 271
Score: 127.00 Matches: 33
Percent Similarity: 50.00% Conservative: 14
Best Local Similarity: 35.11% Mismatches: 37
Query Match: 6.29% Indels: 10
DB: 1 Gaps: 3

US-09-856-061-3 (1-1129) x CRK_DROME (1-271)

QY 841 TCTGATAGAAAGATGTCACACATGATGTCATGATGATACAGCCGCGCAGCA 900
DB 7 Serasparg-----AsnserItyrPheglyPrometSerargInasp 21

QY 901 GTGAGAGAGCCATTCATGAGAGAGAGAGAGATGATGTTCTGTCGAGATGTTCC 960
DB 22 AlathrgluValleuMetAsnGluArgGluArgGlyValPheLeuValArgAspSerAsn 41
QY 961 ACAAAATCCAG 1020
DB 42 SerllealagIasp-----TyValleuysValArgGluAspPrhLysValSerAsn 59
QY 1021 GTAAAAATCCGCTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1080
DB 60 TyrillealasnLysValGlnGlnGlnAspGlnIleValTyrArgIle-----Gly 76
QY 1081 GATGAG 1122
DB 77 AspGlnserPheAspAsnLeuProlLysLeuLeuThrPheTyr 90

Search completed: April 21, 2003, 12:22:40
Job time : 30.2418 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:22:47 ; Search time 30.5993 Seconds
(without alignments)
1637.367 Million cell updates/sec

Title: US-09-856-061-4

Perfect score: 2020

Sequence: 1 FGNFSIPKRNRSWPRIINSATG.....GLRGDEKFDSEDIIEHYKN 376

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2020	100.0	376	22	AA8B1071
2	2009	99.5	428	23	AAE22609
3	2009	99.5	443	23	AAE22608
4	1694	83.9	353	23	AAE22610
5	1132	56.0	435	22	AA8B1070
6	349	17.3	84	21	AA603152
7	289	14.3	449	19	AAW59866
8	289	14.3	456	22	AA893406
9	286	14.2	456	19	AAW59865
10	273	13.3	457	19	AAW59867

11	238	11.8	46	22	AB838147	Peptide #5653 enco
12	238	11.8	46	22	AB823354	Protein #5353 enco
13	238	11.8	46	22	AAW58783	Human brain expres
14	238	11.8	46	22	AAW71296	Human bone marrow
15	238	11.8	46	22	AAW18973	Peptide #5407 enco
16	238	11.8	46	22	AAW31575	Peptide #5612 enco
17	238	11.8	46	23	ABG41096	Human peptide enco
18	236	11.7	95	20	AAW70587	Human SLP-76 SH2 d
19	214	10.6	40	22	AB839614	Peptide #7120 enco
20	214	10.6	40	22	AAW60320	Human bone marrow
21	214	10.6	40	22	AAW72953	Human brain expres
22	214	10.6	40	22	AAW3179	Peptide #7216 enco
23	214	10.6	40	23	ABG42791	Human peptide enco
24	170	8.4	503	22	ABG22854	Novel human diago
25	167	8.3	945	22	AB866077	Drosophila melanog
26	147.5	7.3	759	22	AB821983	Novel human diago
27	147	7.3	724	16	AAW85784	Human GRB-1. Homo
28	147	7.3	724	23	ABG60297	Mouse lymphoma ass
29	146.5	7.3	724	13	AAW26060	Growth Factor Rece
30	146.5	7.3	724	21	AAW97143	Human PI3K p85 pol
31	146.5	7.3	724	22	AAW67442	Amino acid sequenc
32	146.5	7.3	724	22	AAW67621	Human lymphoma ass
33	146.5	7.3	724	23	ABG60298	p85alpha-green flo
34	146.5	7.3	968	19	AAW85011	p85alpha-green flo
35	146.5	7.3	970	19	AAW85020	Human C11A intera
36	144.5	7.2	945	20	AAW06119	Mouse haematopiet
37	143	7.1	659	17	AAW06708	Mouse Bruton's tyr
38	143	7.1	659	23	ABW76488	Mouse Bruton's tyr
39	142	7.0	611	22	ABG04956	Novel human diago
40	140	6.9	635	22	AAU09009	Human Bruton's tyr
41	140	6.9	659	17	AAW94534	BRK tyrosine kinas
42	140	6.9	659	23	ABW76487	Human Bruton's tyr
43	137	6.8	442	17	AAW06709	Human haematopiet
44	136.5	6.8	841	22	ABW67384	Drosophila melanog
45	136.5	6.8	841	22	AAW59217	D.melanogaster cor

ALIGNMENTS

RESULT 1
ID AA8B1071 standard; Protein; 376 AA.
XX
AC AA8B1071.
XX
DT 25-JUN-2001 (first entry)
XX
DE Human mast cell-specific immunoreceptor signal transducer.
XX
KW Mast cell; signal transduction; human; allergic disease; MIST.
KW mast cell-specific immunoreceptor signal transducer.
XX
OS Homo sapiens.
XX
PN JP3146204-B1.
XX
PD 12-MAR-2001.
XX
PF 17-SEP-1999; 99JP-0263778.
XX
PR 17-SEP-1999; 99JP-0263778.
XX
PA (KAGA-) KAGAKU GIUTUSU SHINKO JIGYODAN.
XX
DR WPI; 2001-310022/33.
XX
DR N-PSDB; AAF86140.
XX
PT Mast cell-specific signal-transduction molecule; useful for screening
PT therapeutic compounds for treating allergies; is specifically expressed
PT by mouse mast cell -
XX
PS Example 1; Page 10-11; 12pp; Japanese.

XX This sequence represents human MIST (mast cell-specific immunoreceptor
 CC signal transducer). The invention relates to cDNA encoding a murine mast
 CC cell-specific signal transduction protein. Included in the invention are
 CC cDNA and protein sequences of the mast cell-specific signal transduction
 CC molecule and an expression vector containing the polynucleotide sequence.
 CC The coding sequence of the signal transduction protein can be used for
 CC screening therapeutic compounds which will be useful for treating
 CC allergic diseases.

XX Sequence 376 AA:

Query Match 100.0%; Score 2020; DB 22; Length 376;
 Best Local Similarity 100.0%; Pred. No. 2, 2e-172;
 Matches 376; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 FQNSLPRNRSMPRINSATGQYQRMKPLDMERNFAVLGAGSHDDYDDPELRME 60
 DB 1 FQNSLPRNRSMPRINSATGQYQRMKPLDMERNFAVLGAGSHDDYDDPELRME 60
 QY 61 TWOSIKILPAPRIKESVADTHYFKVAMDTPLDTRTSISIGOPTWNTQRLERVDKPI 120
 DB 61 TWOSIKILPAPRIKESVADTHYFKVAMDTPLDTRTSISIGOPTWNTQRLERVDKPI 120
 QY 121 SRDVSQNIKGDASYRKNKILPPEPRPLITLPRKYQPLPPESSRPLSQRHTEPEVOG 180
 DB 121 SRDVSQNIKGDASYRKNKILPPEPRPLITLPRKYQPLPPESSRPLSQRHTEPEVOG 180
 QY 181 MPSQISLRLDSEVLEAEVPHNQKREESTHLENONTOEIPLAISSSFTSNHVSQNRD 240
 DB 181 MPSQISLRLDSEVLEAEVPHNQKREESTHLENONTOEIPLAISSSFTSNHVSQNRD 240
 QY 241 HRGGMQPCSPQRCOPPASCSPHENILPYKYSWPRPFKRSRDKDVQNHENYIGEYSRQA 300
 DB 241 HRGGMQPCSPQRCOPPASCSPHENILPYKYSWPRPFKRSRDKDVQNHENYIGEYSRQA 300
 QY 301 VEEAFMKENKDGSLVRCSTKSKEEYVLAFFYENKYNKIRPLENQOFALGTGLRG 360
 DB 301 VEEAFMKENKDGSLVRCSTKSKEEYVLAFFYENKYNKIRPLENQOFALGTGLRG 360
 QY 361 DEKPDVEDITEHYKN 376
 DB 361 DEKPDVEDITEHYKN 376

RESULT 2
 AAE22609

ID AAE22609 standard; Protein: 428 AA.

XX AAE22609;

DT 26-JUL-2002 (first entry)

DE Human MIST splice variant protein from clone #7.

KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
 KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
 KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.

OS Homo sapiens.

PN WO200226986-A2.

PD 04-APR-2002.

PF 28-SEP-2001; 2001WO-US30593.

PR 29-SEP-2000; 2000US-237030P.

PA (BRIM) BRISTOL-MYERS SQUIBB CO.

PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB.

XX

DR WPI: 2002-372126/40.

DR N-PSDB: AAD35801.

XX New isolated mast cell immunoreceptor signal transducer polypeptide,

PT useful for treating immune disorder involving hyperactivity of B- or

PS T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma

XX Claim 11; Fig 5; 174pp; English.

CC The present invention relates to novel mast cell immunoreceptor signal
 CC transducer (MIST) proteins and polynucleotides encoding such proteins.
 CC MIST sequences of the invention are useful for preventing, treating or
 CC ameliorating a medical condition in mammalian subject. They are useful
 CC for treating an immune disorder involving hyperactivity of B- or T-
 CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
 CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
 CC as targets for therapeutic intervention in immune cell disorders and
 CC inflammatory indications, for diagnosis and/or screening of disorders
 CC or diseases associated with expression of MIST, for screening for
 CC antagonists or inhibitors of the interaction of MIST with cellular
 CC signalling components. They are used in assays that detect activation
 CC or induction of various B and T-cell-related neoplasms or cancers.
 CC Sequences of the invention are also used in gene therapy. The present
 CC sequence is human MIST splice variant protein from clone #7.

SQ Sequence 428 AA:

Query Match 99.5%; Score 2009; DB 23; Length 428;
 Best Local Similarity 99.5%; Pred. No. 2, 6e-171;
 Matches 374; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 FQNSLPRNRSMPRINSATGQYQRMKPLDMERNFAVLGAGSHDDYDDPELRME 60
 DB 19 FQNSLPRNRSMPRINSATGQYQRMKPLDMERNFAVLGAGSHDDYDDPELRME 78
 QY 61 TWOSIKILPAPRIKESVADTHYFKVAMDTPLDTRTSISIGOPTWNTQRLERVDKPI 120
 DB 79 TWOSIKILPAPRIKESVADTHYFKVAMDTPLDTRTSISIGOPTWNTQRLERVDKPI 138
 QY 121 SRDVSQNIKGDASYRKNKILPPEPRPLITLPRKYQPLPPESSRPLSQRHTEPEVOG 180
 DB 139 SRDVSQNIKGDASYRKNKILPPEPRPLITLPRKYQPLPPESSRPLSQRHTEPEVOG 198
 QY 181 MPSQISLRLDSEVLEAEVPHNQKREESTHLENONTOEIPLAISSSFTSNHVSQNRD 240
 DB 199 MPSQISLRLDSEVLEAEVPHNQKREESTHLENONTOEIPLAISSSFTSNHVSQNRD 258
 QY 241 HRGGMQPCSPQRCOPPASCSPHENILPYKYSWPRPFKRSRDKDVQNHENYIGEYSRQA 300
 DB 259 HRGGMQPCSPQRCOPPASCSPHENILPYKYSWPRPFKRSRDKDVQNHENYIGEYSRQA 318
 QY 301 VEEAFMKENKDGSLVRCSTKSKEEYVLAFFYENKYNKIRPLENQOFALGTGLRG 360
 DB 319 VEEAFMKENKDGSLVRCSTKSKEEYVLAFFYENKYNKIRPLENQOFALGTGLRG 378
 QY 361 DEKPDVEDITEHYKN 376
 DB 379 DEKPDVEDITEHYKN 394

RESULT 3
 AAE22608

ID AAE22608 standard; Protein: 443 AA.

XX AAE22608;

DT 26-JUL-2002 (first entry)

DE Human MIST protein #1.

KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
 KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
 KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.

```

XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX FT Modified-site 84
XX FT Modified-site /label= Tyrosine_phosphorylation_binding_site
XX FT Modified-site 111
XX FT Modified-site /label= Tyrosine_phosphorylation_binding_site
XX FT Domain 306..311
XX FT Domain /label= SH3-binding_proline-rich_motif
XX FT Domain 324..407
XX FT Domain /label= SH2_domain
XX PN WO200226986-A2.
XX PD 04-APR-2002.
XX PF 28-SEP-2001; 2001WO-US30593.
XX PR 29-SEP-2000; 2000US-237030P.
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX DR WPI: 2002-372126/40.
XX DR N-PSDB: AAD35800.
XX PS New isolated mast cell immunoreceptor signal transducer polypeptide,
XX PT useful for treating immune disorder involving hyperactivity of B- or
XX PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
XX PS Claim 11; Fig 3; 171pp; English.
XX CC The present invention relates to novel mast cell immunoreceptor signal
XX CC transducer (MIST) proteins and polynucleotides encoding such proteins.
XX CC MIST sequences of the invention are useful for preventing, treating or
XX CC ameliorating a medical condition in mammalian subject. They are useful
XX CC for treating an immune disorder involving hyperactivity of B- or T-
XX CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX CC lymphoma, tumor or thymoma in a mammal. MIST sequences are useful
XX CC as targets for therapeutic intervention in immune cell disorders and
XX CC inflammatory indications, for diagnosis and/or screening of disorders
XX CC or diseases associated with expression of MIST, for screening for
XX CC antagonists or inhibitors of the interaction of MIST with cellular
XX CC signaling components. They are used in assays that detect activation
XX CC or induction of various B and T-cell-related neoplasms or cancers.
XX CC Sequences of the invention are also used in gene therapy. The present
XX CC sequence is human MIST protein.
XX SQ Sequence 443 AA;
XX Query Match 99.5%; Score 2009; DB 23; Length 443;
XX Best Local Similarity 99.5%; Pred. No. 2.7e-171;
XX Matches 374; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

DB 274 HRCGMPCSPQRCOPRASCSPHENILPYKYSWRPPPKRSRDKVDYHNEMWIGEYSRQA 333
QY 301 VEEAFMKENKDGSEFLVRCSTKSEEPYVLAVREKNVYVWKTRFLERNOQFALGTGLRG 360
DB 334 VEEAFMKENKDGSEFLVRCSTKSEEPYVLAVREKNVYVWKTRFLERNOQFALGTGLRG 393
QY 361 DEKFDSEVEDIEEHYKN 376
DB 394 DEKFDSEVEDIEEHYKN 409
RESULT 4
AAE22610
ID AAE22610 standard; Protein: 353 AA.
XX AC AAE22610;
XX AC AAE22610;
XX DT 26-JUL-2002 (first entry)
XX DE Human MIST splice variant protein from clone #12.
XX KW Human; mast cell immunoreceptor signal transducer; MIST; immune disorder;
XX KW lymphocyte; lymphoma; tumour; thymoma; immune cell disorder; neoplasm;
XX KW inflammation; cancer; gene therapy; cytostatic; immunomodulatory.
XX OS Homo sapiens.
XX PN WO200226986-A2.
XX PD 04-APR-2002.
XX PF 28-SEP-2001; 2001WO-US30593.
XX PR 29-SEP-2000; 2000US-237030P.
XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX PI Perez-Villar JJ, Chang H, Yang W, Wu Y, Whitney GS, Kanner SB;
XX DR WPI: 2002-372126/40.
XX DR N-PSDB: AAD35802.
XX PS New isolated mast cell immunoreceptor signal transducer polypeptide,
XX PT useful for treating immune disorder involving hyperactivity of B- or
XX PT T-lymphocytes, and for inhibiting T-lymphoma, tumor or thymoma
XX PS Claim 11; Fig 8; 171pp; English.
XX CC The present invention relates to novel mast cell immunoreceptor signal
XX CC transducer (MIST) proteins and polynucleotides encoding such proteins.
XX CC MIST sequences of the invention are useful for preventing, treating or
XX CC ameliorating a medical condition in mammalian subject. They are useful
XX CC for treating an immune disorder involving hyperactivity of B- or T-
XX CC lymphocytes in a mammal, for inhibiting growth of or inhibiting T-
XX CC lymphoma, tumour or thymoma in a mammal. MIST sequences are useful
XX CC as targets for therapeutic intervention in immune cell disorders and
XX CC inflammatory indications, for diagnosis and/or screening of disorders
XX CC or diseases associated with expression of MIST, for screening for
XX CC antagonists or inhibitors of the interaction of MIST with cellular
XX CC signaling components. They are used in assays that detect activation
XX CC or induction of various B and T-cell-related neoplasms or cancers.
XX CC Sequences of the invention are also used in gene therapy. The present
XX CC sequence is human MIST splice variant protein from clone #12.
XX SQ Sequence 353 AA;
XX Query Match 83.9%; Score 1694; DB 23; Length 353;
XX Best Local Similarity 99.4%; Pred. No. 3.2e-143;
XX Matches 317; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

Db 1 MEETWOSIKILPAPRIKESVADTHYFKVAMDPLPLDTRTISIGOPTWNTQTRLERYD 60
 QY 118 KPISRDVRSONIKGDASVYKNNKIPPLPPRLITLPPKKYQPLPEPSSRPPLSORHTPE 177
 Db 61 KPISRDVRSONIKGDASVYKNNKIPPLPPRLITLPPKKYQPLPEPSSRPPLSORHTPE 120
 QY 178 VQGMPSQISLRDLSEVLEAEKYPHNQKRPSTHLENNQTOEIPLAISSSFTTSHSVQ 237
 Db 121 VQGMPSQISLRDLSEVLEAEKYPHNQKRPSTHLENNQTOEIPLAISSSFTTSHSVQ 180
 QY 238 NBDHGGMOPCSPORCOPPASCSPHENILPYKTYTSWRPFPRKSDKDYQHNEMTIGES 297
 Db 181 NBDHGGMOPCSPORCOPPASCSPHENILPYKTYTSWRPFPRKSDKDYQHNEMTIGES 240
 QY 238 RQAVEAEAFMKENKDGSEFLVRCSTSKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 357
 Db 241 RQAVEAEAFMKENKDGSEFLVRCSTSKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 300
 QY 358 LRGEKFDSEVEDIIEHYKN 376
 Db 301 LRGEKFDSEVEDIIEHYKN 319

RESULT 5

AAB81070
 ID AAB81070 standard; Protein: 435 AA.

XX AAB81070;

DT 25-JUN-2001 (first entry)

XX DE Murine mast cell-specific signal transduction protein.

XX KW Mast cell; signal transduction; mouse; allergic disease.

XX OS Mus musculus.

XX PN JP3146204-B1.

XX PD 12-MAR-2001.

XX PE 17-SEP-1999; 99JP-0263778.

XX PR 17-SEP-1999; 99JP-0263778.

XX PA (KAGA-) KAGAKU GIJUTSU SHINKO JIGYODAN.

XX DR WPI: 2001-310022/33.

XX DR N-PSDB: AAF86139.

XX PT Mast cell-specific signal-transduction molecule, useful for screening
 PT therapeutic compounds for treating allergies, is specifically expressed
 PT by mouse mast cell -

XX PS Claim 1: Page 7-8; 12pp; Japanese.

XX CC This sequence represents a murine mast cell-specific signal transduction
 CC protein. The invention includes the cDNA and protein sequences of the
 CC mast cell-specific signal transduction molecule and an expression vector
 CC containing the polynucleotide sequence. The coding sequence of the signal
 CC transduction protein can be used for screening therapeutic compounds
 CC which will be useful for treating allergic diseases.

XX SQ Sequence 435 AA;

Query Match 56.0%; Score 1132; DB 22; Length 435;

Best Local Similarity 60.9%; Pred. No. 9.6e-93;

Matches 229; Conservative 42; Mismatches 101; Indels 4; Gaps 3;

QY 1 FONSILPKNRSMRPNSTGQYQYKNNKRLDMMERFAVLDGAGHSDDDVDDELKMEF 60
 Db 19 FONSILPKNRSMRPNSTGQYQYKNNKRLDMMERFAVLDGAGHSDDDVDDELKMEF 60

QY 61 TWOSIKILPAPRIKESVADTHYFKVAMDPLPLDTRTISIGOPTWNTQTRLERYD 119
 Db 79 AMPSKILPAPRIKESVADTHYFKVAMDPLPLDTRTISIGOPTWNTQTRLERYD 138
 QY 120 ISRDVRSONIKGDASVYKNNKIPPLPPRLITLPPKKYQPLPEPSSRPPLSORHTPE 178
 Db 139 TEKDVRSQFQKFKYTKIKNKTPLPPRPAITLPPKKYQPLPEPSSRPPLSORHTPE 198
 QY 179 QGMPSQISLRDLSEVLEAEKYPHNQKRPSTHLENNQTOEIPLAISSSFTTSHSVQ 238
 Db 199 QGMPSQISLRDLSEVLEAEKYPHNQKRPSTHLENNQTOEIPLAISSSFTTSHSVQ 258
 QY 233 RQAVEAEAFMKENKDGSEFLVRCSTSKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 298
 Db 259 RQAVEAEAFMKENKDGSEFLVRCSTSKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 316
 QY 299 QAVEAEAFMKENKDGSEFLVRCSTSKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 358
 Db 317 QAVEAEAFMKENKDGSEFLVRCSTSKSEEPYLAIFYENKYNVVKIRFLERNOQFALGTG 376
 QY 359 RGDEKFDSEVEDIIEHY 374
 Db 377 RGDEKFDSEVEDIIEHY 392

RESULT 6

AAG03152
 ID AAG03152 standard; Protein: 84 AA.

XX AAG03152;

DT 06-OCT-2000 (first entry)

XX DE Human secreted protein, SEQ ID NO: 7233.

XX KW Human; 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
 KW gene therapy; chromosome mapping.

XX OS Homo sapiens.

XX PN EP1033401-A2.

XX PD 06-SEP-2000.

XX PE 21-FEB-2000; 2000EP-0200610.

XX PR 26-FEB-1999; 99US-0122487.

XX PA (GEST) GENSET.

XX PI Dunas Milne Edwards J, Duclert A, Giordano J;

XX DR WPI: 2000-500381/45.

XX DR N-PSDB: AAC03158.

XX PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
 PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
 PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
 XX Claim 13: SEQ ID 7233; 71pp + CD-ROM; English.

XX CC The present sequence is a polypeptide encoded by one of a large number
 CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
 CC were prepared from total human RNAs or polyA+ RNAs derived from 30
 CC different tissues. EST sequences usually correspond mainly to the 3'
 CC untranslated region (UTR) of the mRNA because they are often obtained
 CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
 CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
 CC those cases where longer cDNA sequences have been obtained, the full 5'
 CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
 CC ends and can therefore be used to obtain full length cDNAs and genomic
 CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
 CC chromosome mapping procedures. They are used to obtain upstream

CC regulatory sequences and to design expression and secretion vectors.
 XX Sequence 84 AA:
 SQ
 Query Match 17.3%; Score 349; DB 21; Length 84;
 Best Local Similarity 97.0%; Pred. No. 1.3e-23;
 Matches 64; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 FQNFSLPKRMSRPRINSATGQYORNMKPLLDMEERFAVLGAKGSHDDYDPELRME 60
 DB 19 FQNFSLPKRMSRPRINSATGQYORNMKPLLDMEERFAVLGAKGSHDDYDPELRME 78
 OY 61 TWOSIK 66
 DB 79 TWOSIK 84

RESULT 7
 AAM59866 standard; Protein: 449 AA.
 ID AAM59866
 XX AAM59866;
 AC
 XX
 XX
 DT 20-NOV-1998 (first entry)
 XX
 DE Amino acid sequence of the human BLNK-2 protein.
 XX Human; BLNK-2; B cells linker protein-2; apoptosis; Grb2; PLC-gamma;
 KW SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; SOS;
 KW ras pathway; GDP; GTP; calcium pathway; antigen.
 XX Homo sapiens.
 OS
 XX
 XX WO9832852-A1.
 XX
 PD 30-JUL-1998.
 XX
 PF 23-JAN-1998; 98WO-US01394.
 XX
 PR 17-MAR-1997; 97US-0819013.
 PR 24-JAN-1997; 97US-0788322.
 XX
 XX (UNITW) UNIV WASHINGTON.
 PA
 XX Chan AC, Fu C;
 XX WPI; 1998-427948/36.
 DR N-PSDB; AAV41902.
 XX
 XX Human B cell linker proteins - useful in the treatment of diseases
 PT involving increased or decreased apoptosis
 PS Disclosure; Fig 3; 56pp; English.
 XX
 CC This is the amino acid sequence of the human BLNK-2 (B cells linker
 CC protein-2) protein, used the treatment of diseases involving the
 CC increase and decrease of apoptosis, in the method of the invention.
 CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
 CC intracellular calcium levels and Grb2 is an adapter molecule
 CC containing two SH3 domains that mediate its interaction with the
 CC guanine nucleotide exchange factor. Son of Sevenless (SOS) which in
 CC turn activates the ras pathway by facilitating the exchange of GDP for
 CC GTP on the ras molecule. Activation of both ras and calcium pathways
 CC are required for efficient B cell antigen receptor function. BLNK
 CC binding proteins can be used to identify BLNK proteins in a target
 CC sample.
 XX
 XX Sequence 449 AA:
 Query Match 14.3%; Score 289; DB 19; Length 449;
 Best Local Similarity 25.6%; Pred. No. 3.3e-17;
 Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;

OY 43 AKGSHDDYDPELRMEETWOSIKILPAPRIKSEYADTHYFKVAMDTPLDTRTSISI 102
 DB 81 AENMADDSYEPPEVE-----TRPVLPALPANGER-----IDNRSS----- 118
 OY 103 GQFTWNTQTRLEKVRDKPDRVRSNLIKGDASYRKKKILPLPPRLITLKKYQPLPEP 162
 DB 119 -----QHSPPFSKTLPSK-----PSWSEKARLTSLPALALQKPQ-VPPKP 161
 OY 163 -----ESSRP-----FLSQRHTTPEVQ-----GMSQSLNRD 189
 DB 162 KGLLEDEADYVPEVDENDENYIHTPESSSPPEKAPVNVNSTPNSSTPASPGTASGRN 221
 OY 190 LSEVLPAEK-----VPHNORP-----ESTHLENONTQELPLAISSSFTSNHSVQ 237
 DB 222 -SGAMETKSPPPAAPSPPLPACGKKPTTTLKTTVYASQD-----ASSVCEKRP 270
 OY 238 NRDRGG--WQPCSPQRCOPPASCSPHENILPY-KYT-SWRP---PFPKRSRDKVOHN 289
 DB 271 AERHRGSSHRQEAIVQSPVFPFAQKQIHQKPIPLRFTEGNPVVDGFLPFSSSNSTISEQ 330
 OY 290 E-----WTIGESRQAVAEAFMKENKDGSEFLVRODCTSKSEEPYLVAFYENKYNK 342
 DB 331 EAGVLCKPWTYAGACDRKSEALHRSNKDSEFLIRSSGHSKQPYTLVVFENKRYINP 390
 OY 343 IRFLERNQOPALGTGLRGDEKFDSEVDIEHYKN 376
 DB 391 VRFEATKQYALGRKKNGEYFGSVAETIRNHQ 424

RESULT 8
 AAB93406 standard; Protein: 456 AA.
 ID AAB93406
 XX AAB93406;
 AC
 XX
 XX
 DT 26-JUN-2001 (first entry)
 XX
 DE Human protein sequence SEQ ID NO:12602.
 XX
 XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 KW Homo sapiens.
 OS
 XX
 XX EP1074617-A2.
 XX
 PD 07-FEB-2001.
 XX
 PF 28-JUL-2000; 2000EP-0116126.
 XX
 PR 29-JUL-1999; 99JP-0248036.
 PR 27-AUG-1999; 99JP-0300253.
 PR 11-JAN-2000; 2000JP-0118776.
 PR 02-MAY-2000; 2000JP-0183767.
 PR 09-JUN-2000; 2000JP-0241899.
 XX
 XX (HELT-) HELIX RES INST.
 PA
 XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX WPI; 2001-318749/34.
 DR
 XX
 PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 PS
 XX Claim 8; SEQ ID 12602; 2537pp + CD ROM; English.
 XX
 CC The present invention describes primer sets for synthesizing 5602
 CC full-length cDNAs defined in the specification. Where a primer set
 CC comprises: (a) an oligo-dt primer and an oligonucleotide complementary
 CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
 CC of an oligonucleotide comprising a sequence complementary to the
 CC complementary strand of a polynucleotide which comprises a 5'-end
 CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in
 CC the specification. The primer sets can be used in antisense therapy and
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to
 CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632
 CC represent oligonucleotides, all of which are used in the exemplification
 CC of the present invention.

XX Sequence 456 AA;

Query Match 14.3%; Score 289; DB 22; Length 456;

Best Local Similarity 25.6%; Pred. No. 3.4e-17;
 Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;

QY 43 AKGSHDDYDDELMEETWOSIKILPAPRIKESYADTHYKAMDTPLDPTTSISI 102
 DB 88 AENADDSYEPPEVQE-----TRVHPALPARGFY-----IDNRSS--- 125
 QY 103 GQPTNNTQRLERVDKPSRVSQNIKGDASVRKNKILPLPPRLITLPKKYQPLPEP 162
 DB 126 -----QRHSPPSKTLPSK-----PSWPEKARLTSTLPALTALQKPQ--VPPKP 168
 QY 163 -----ESSRP-----PLSORHTPEVQ--GMPQSISLRD 189
 DB 169 KGLEDEADYVVPEDNDENTIHPTSSSPPEKAPMNRKSTKPSNPASPGTASGRN 228
 QY 190 LSEVLEAEK-----VPHNOKRP-----ESTHLENONTOEILPLAISSFTTSNHSVQ 237
 DB 229 -SGAMETKSPPAAPSPPLPRAGKKPTPLKTTVPASQON-----ASSVCEKEPIP 277
 QY 228 NRDRHGG--MPCSPQRCQPPASCSPHENILPY-KYT-SWRP-----PPKASDKRDQVHN 289
 DB 278 AERHGGSHRQDAVOSVPEPPAQKQIHPKPLPRFTGEGNPTVDGPIPLFSSNSTISEQ 337
 QY 290 E-----WYIGEYSRQAVEAEAFMKENKDGSLVRDCTSKSEPPYVLAFFENKYNVVK 342
 DB 338 EAGVLCKPWTAGACDRKSAEALHRSNKGDSFLIRKSSGHSKQPYTLVFFNKRYNIP 397
 QY 343 IRLERNOQFALGTGLRDEKFDVSDIEDIIEHYKN 376
 DB 398 VRFTEATKQYALGRKKNGEEYFGSVAETIRNHQH 431

RESULT 9

AAW59865

ID AAW59865 standard; Protein; 456 AA.

AC AAW59865;

DT 20-NOV-1998 (first entry)

XX Amino acid sequence of the human BLNK-1 protein.

XX Human; BLNK-1; B cells linker protein-1; apoptosis; Grb2; PLC-gamma;

KW SH domain; guanine nucleotide exchange factor; Son of Sevenless; Sos;

KM ras pathway; GDP; GTP; calcium pathway; antigen.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Misc-difference 327 /note="encoded by AGC"

FT

XX W09832852-A1.

XX 30-JUL-1998.

XX 23-JAN-1998; 98MO-US01394.

XX 17-MAR-1997; 97US-0819013.

XX 24-JAN-1997; 97US-0788322.

XX (UNIT) UNIV WASHINGTON.

XX Chan AC, Fu C;

XX WPI. 1998-427948/36.

XX N-PSDB; AAW41901.

XX Human B cell linker proteins - useful in the treatment of diseases

XX involving increased or decreased apoptosis

XX Claim 4; Fig 1; 56pp; English.

XX This is the amino acid sequence of the human BLNK-1 (B cells linker

XX protein-1) protein, used the treatment of diseases involving the

XX increase and decrease of apoptosis, in the method of the invention.

XX BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates

XX intracellular calcium levels and Grb2 is an adapter molecule

XX containing two SH3 domains that mediate its interaction with the

XX guanine nucleotide exchange factor, Son of Sevenless (Sos) which in

XX turn activates the ras pathway by facilitating the exchange of GDP for

XX GTP on the ras molecule. Activation of both ras and calcium pathways

XX are required for efficient B cell antigen receptor function. BLNK

XX binding proteins can be used to identify BLNK proteins in a target

XX sample.

XX Sequence 456 AA;

Query Match 14.2%; Score 286; DB 19; Length 456;

Best Local Similarity 25.6%; Pred. No. 6.2e-17;
 Matches 101; Conservative 51; Mismatches 132; Indels 110; Gaps 17;

QY 43 AKGSHDDYDDELMEETWOSIKILPAPRIKESYADTHYKAMDTPLDPTTSISI 102
 DB 88 AENADDSYEPPEVQE-----TRVHPALPARGFY-----IDNRSS--- 125
 QY 103 GQPTNNTQRLERVDKPSRVSQNIKGDASVRKNKILPLPPRLITLPKKYQPLPEP 162
 DB 126 -----QRHSPPSKTLPSK-----PSWPEKARLTSTLPALTALQKPQ--VPPKP 168
 QY 163 -----ESSRP-----PLSORHTPEVQ--GMPQSISLRD 189
 DB 169 KGLEDEADYVVPEDNDENTIHPTSSSPPEKAPMNRKSTKPSNPASPGTASGRN 228
 QY 190 LSEVLEAEK-----VPHNOKRP-----ESTHLENONTOEILPLAISSFTTSNHSVQ 237
 DB 229 -SGAMETKSPPAAPSPPLPRAGKKPTPLKTTVPASQON-----ASSVCEKEPIP 277
 QY 228 NRDRHGG--MPCSPQRCQPPASCSPHENILPY-KYT-SWRP-----PPKASDKRDQVHN 289
 DB 278 AERHGGSHRQDAVOSVPEPPAQKQIHPKPLPRFTGEGNPTVDGPIPLFSSNSTISEQ 337
 QY 290 E-----WYIGEYSRQAVEAEAFMKENKDGSLVRDCTSKSEPPYVLAFFENKYNVVK 342
 DB 338 EAGVLCKPWTAGACDRKSAEALHRSNKGDSFLIRKSSGHSKQPYTLVFFNKRYNIP 397
 QY 343 IRLERNOQFALGTGLRDEKFDVSDIEDIIEHYKN 376
 DB 398 VRFTEATKQYALGRKKNGEEYFGSVAETIRNHQH 431

RESULT 10

AAW59867

ID AAW59867 standard; Protein; 457 AA.


```

XX AC AAM59867;
XX 20-NOV-1998 (first entry)
XX DE Amino acid sequence of the mouse BLNK protein.
XX KW Mouse; BLNK; B cells linker protein; apoptosis; Grb2; PLC-gamma;
XX KM SH3 domain; guanine nucleotide exchange factor; Son of Sevenless; SOS;
XX KX ras pathway; GDP; GTP; calcium pathway; antigen.
XX OS Mus sp.
XX PN W09832852-A1.
XX PD 30-JUL-1998.
XX PF 23-JAN-1998; 98MO-US01394.
XX PR 17-MAR-1997; 97US-0819013.
XX PR 24-JAN-1997; 97US-0788322.
XX PA (UNITW ) UNIV WASHINGTON.
XX PI Chan AC, Fu C;
XX DR WPI; 1998-427948/36.
XX DR N-PSDB; AAV41903.
XX PT Human B cell linker proteins - useful in the treatment of diseases
XX PT involving increased or decreased apoptosis
XX PS Disclosure; Fig 5; 56pp; English.
XX XX This is the amino acid sequence of the murine BLNK (B cells linker
CC protein) protein, used the treatment of diseases involving the
CC increase and decrease of apoptosis. In the method of the invention.
CC BLNK proteins interact with Grb2 and PLC-gamma. PLC-gamma regulates
CC intracellular calcium levels and Grb2 is an adapter molecule
CC containing two SH3 domains that mediate its interaction with the
CC guanine nucleotide exchange factor, Son of Sevenless (SOS) which in
CC turn activates the ras pathway by facilitating the exchange of GDP for
CC GTP on the ras molecule. Activation of both ras and calcium pathways
CC are required for efficient B cell antigen receptor function. BLNK
CC binding proteins can be used to identify BLNK proteins in a target
CC sample.
CC XX
SQ Sequence 457 AA;
Query Match 13.5%; Score 273; DB 19; Length 457;
Best Local Similarity 23.8%; Pred. No. 9,1e-16;
Matches 96; Conservative 60; Mismatches 155; Indels 92; Gaps 16;
QY 32 WERFAAVALGAKGHS-----DDDDPELRMEETWQSIKILPAPRIKESY 78
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 64 WSDDFSDYENPDHSDSEMYVPAEETGDSYEPPEAQ-----QTVVHPALFTTGEY 119
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 79 ADTHYFVAMDPLPLDTRISISIGOPTWNTQTRLEVDKPFISDVSONIKGASVRYKN 138
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 120 VDNK-----SSQRHSPFESKPLPS--KPSW-----PSAKARLASTLPADNSLOKP 162
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 139 KILPPEPRPLI-----TLPKKYQPLPEP-----ESSRPLSQRHFP 176
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 163 QVP-PKRDLEDEADYVPEVDENDENTHPRSSPPAEKAPVWNSTKFNSSSKHMS 221
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 177 --EVQAMPQSILRDLESEVLEAEKVPVHQRKESTHLENTQEIPLATSSSFTTNNH 234
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 222 PGTVAGRNISGVWDSKSLPAPSPPLPAGKKP-ATPL-----KTTVPV-PLPNASVCEK 275
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 235 SVQNRDHRGG--MPCSFQRCQPPASCSPHENILPYKTSRPPPPKRA----- 280
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 276 FVPAERHRGSSHRDVTQSPVFPPTOKPVHOKPVPL-----PRFPEAGSPADGPFHSF 329
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |

```

```

QY 281 -----SDRK-DYQHNENYIGETSRQAVFEAFMKENKSGFLVRCSTSKSKEEPLYAVE 333
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 330 PFNLTFADQDEGLCKPWPYAGACDRKFAEALHRSKDSFLIRSFSGHDSKOPTVLAVF 389
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
QY 334 YENKYVYKIRFLERNQOFALGTGLRGDEKFDSDVEDIIEHYNN 376
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 390 FNRKYNIPIRYFTEATKOYALCKKKNGEYFSGVVEIVNSHOH 432
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |

RESULT 11
ABB38147
ID ABB38147 standard; Peptide: 46 AA.
XX AC ABB38147;
XX DT 04-FEB-2002 (first entry)
XX DE Peptide #5653 encoded by human foetal liver single exon probe.
XX KM Human; foetal liver; gene expression; single exon nucleic acid probe.
XX OS Homo sapiens.
XX OS WO200157277-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001MO-US00669.
XX PR 04-FEB-2000; 2000US-0180312.
XX PR 26-MAY-2000; 2000US-0207456.
XX PR 30-JUN-2000; 2000US-0608408.
XX PR 03-AUG-2000; 2000US-0632366.
XX PR 21-SEP-2000; 2000US-0234687.
XX PR 27-SEP-2000; 2000US-0236359.
XX PR 04-OCT-2000; 2000GB-0024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR WPI; 2001-483447/52.
XX PT Human genome-derived single exon nucleic acid probes useful for
XX PT analyzing gene expression in human fetal liver -
XX PS Claim 27; SEQ ID NO 30782; 639pp + sequence listing; English.
XX XX The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
CC XX
SQ Sequence 46 AA;
Query Match 11.8%; Score 238; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4,8e-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 DGSFLVRCSTSKSKEEPLYAVFTEKNYVYKIRFLERNQOFALGT 356
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |
DB 1 DGSFLVRCSTSKSKEEPLYAVFTEKNYVYKIRFLERNQOFALGT 46
| : : : | : : | : : | : : | : : | : : | : : | : : | : : |

RESULT 12
ABB23354
ID ABB23354 standard; Protein: 46 AA.
XX

```

```

KW epilepsy; cancer.
XX
OS Homo sapiens.
XX
PN WO200157275-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00667.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
DT Single exon nucleic acid probes for analyzing gene expression in human
XX brains -
XX
PS Example 4; SEQ ID NO: 30888; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is a protein encoded by one of
XX the probes of the invention.
XX
SQ Sequence 46 AA:
XX
Query Match 11.8%; Score 238; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4, Be-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0.
QY 311 DGSFLVRDCSTKSKKEEPLYLAVFENKYNVVKIRFLERNQGFALGT 356
DB 1 DGSFLVRDCSTKSKKEEPLYLAVFENKYNVVKIRFLERNQGFALGT 46
RESULT 14
AAAM71296
ID AM71296 standard; Protein; 46 AA.
AC AAM71296;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 31602.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe;
XX microarray; cancer; leukaemia; lymphoma; myeloma.
XX
OS Homo sapiens.
XX
PN WO200157276-A2.
XX
PD 30-JAN-2001; 2001WO-US00668.
XX
PE 09-AUG-2001.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.

```

```

PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 31602; 658bp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 46 AA;
Query Match 11.8%; Score 238; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 DGSFLVRCSTSKSKEEPLYLAFFYENKYNVKIRFLERNQGFALGT 356
DB 1 DGSFLVRCSTSKSKEEPLYLAFFYENKYNVKIRFLERNQGFALGT 46

RESULT 15
AAM18973
ID AAM18973 standard; Protein; 46 AA.
XX
AC AAM18973;
XX
DF 12-OCT-2001 (first entry)
XX
DE Peptide #5407 encoded by probe for measuring cervical gene expression.
XX
KW Probe: human; microarray; gene expression; cervical epithelial cell;
XX
KW cervical cancer.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PE 30-JAN-2001; 2001WO-US00670.
XX
PR 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488901/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human cervical epithelial cells -
XX
PS Claim 27; SEQ ID NO 23799; 487bp; English.
XX

```

```

CC The present invention relates to human single exon nucleic acid probes
CC (SNP: see AAI10068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SNPs are derived from human Hela cells. The SNPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 46 AA;
Query Match 11.8%; Score 238; DB 22; Length 46;
Best Local Similarity 100.0%; Pred. No. 4.8e-14;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 311 DGSFLVRCSTSKSKEEPLYLAFFYENKYNVKIRFLERNQGFALGT 356
DB 1 DGSFLVRCSTSKSKEEPLYLAFFYENKYNVKIRFLERNQGFALGT 46

```

Search completed: April 21, 2003, 12:36:25
Job time : 31.5993 secs


```

Db 88 AEENADDSYEPPEVEOE-----TRPVHPLPFAARGEY-----IDNRSS----- 125
QY 103 GQPTNNTQRLERVDKPIRSRDVRSQNIKGDASVRKKNKIPLPRLPRLITPKKYQPLPPPP 162
Db 126 -----ORHSPPEKTLPSK-----PSWPEKARLTSTLALALQPPQ--VPPRP 168
QY 163 -----ESSRP-----PLSQRTTFPEVQ--GMPQSILSD 189
Db 169 KGLLEADADVVPVEDNDENYIHPRESSSPPEKAPMVNRSTKPNSSSTASPGTASGHN 228
QY 190 LSEVLEAEK-----VPHNORKP-----ESTHLENTQTOIPLAISSSFTTNSHYQ 237
Db 229 -SGAMETSSPPAPSPPLPRAKKKPTPLKTPVASSQN-----ASSVCEKPTP 277
QY 238 NRDHRGCG--MPCSPORCOPASCSPHENILPY-KYT-SWRP-----PPPKRSRDRKDYQHN 289
Db 278 ARRHGSSHROEAVOSPVPFPKOKIHOPIPLPRFTBEGNTTVGCPDIFSSNSTISIQ 337
QY 290 E-----WYIGYSRQAVEAEAFMKENKDGSLVDCSTKSKPEEYVLAVFENKRYNKK 342
Db 338 EAGVLCKRPYAGACDRKSAEALHRSNKGDSFLIRKSSHDCKOPYTLVVFNNKRYNIP 397
QY 343 IRFLERNQOAFALGTGRGDEKDSVEDIEIHYKN 376
Db 398 VRIETATKOYALGRKKRNGEYEGSVAEIIRNHQ 431

```

RESULT 2

```

US-07-906-349A-5
; Sequence 5, Application US/07906349A
; Patent No. 3434064

```

```

; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE KINASES AN
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Browdy and Neimark
; STREET: 419 Seventh Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/906,349A
; FILING DATE: 30-JUN-1992
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/643,237
; FILING DATE: 18-JAN-1991
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 724 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-906-349A-5

```

```

Query Match 7.38; Score 147; DB 1; Length 724;
Best Local Similarity 24.78; Pred. No. 2,1e-05;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;

```

```

QY 176 PEVQGMPSQISLRDISEVLEAEKVPHNQ-----RKDESTHLENTQTOEI- 220
Db 211 PEVQSSSEYIQL--LKKLIRSPSIPHOYWLTLQYLKHFELKSQSSKNLLNARLSEIF 268
QY 221 -PLAISSSFTTNSH-----SVONRDRHGMQPCSPORCOPASCSPHENILPY 268
Db 269 SPMLFRFSAASDNTNENIKVIEILISTEMNERQ-----PAPALPP----- 309
QY 269 KITSNRPPPK-----RSRDRQVHNEMYIGYSRQAVEAEAFMKENKDGSLVDCSTK 322
Db 310 -----KPKKPTTVANNGMNMSLQNAEYWMODISREEVNEK-LRDTADGTFLVDASTK 363
QY 322 SKEEYVLAVFENKRYNKK-IRFLERNQOAFALGTGRGDEKDSVEDIEIHYKN 376
Db 364 MHGD-YTLTL---RKGNKKLIKIFHRDQKYGFSPL---TFSSVELINHYRN 410

```

RESULT 3

```

US-08-167-035-2
; Sequence 2, Application US/08167035
; Patent No. 5618691

```

```

; GENERAL INFORMATION:
; APPLICANT: Schlessinger, Joseph
; APPLICANT: Skolnick, Edward Y.
; APPLICANT: Margolis, Benjamin L.
; TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
; TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: New York
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/167,035
; FILING DATE: 16-DEC-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-062
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8664
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 724 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
US-08-167-035-2

```

```

Query Match 7.38; Score 147; DB 1; Length 724;
Best Local Similarity 24.78; Pred. No. 2,1e-05;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;
QY 176 PEVQGMPSQISLRDISEVLEAEKVPHNQ-----RKDESTHLENTQTOEI- 220
Db 211 PEVQSSSEYIQL--LKKLIRSPSIPHOYWLTLQYLKHFELKSQSSKNLLNARLSEIF 268
QY 221 -PLAISSSFTTNSH-----SVONRDRHGMQPCSPORCOPASCSPHENILPY 268
Db 269 SPMLFRFSAASDNTNENIKVIEILISTEMNERQ-----PAPALPP----- 309

```

```

0y 323 SKPEPYLVAVEEKKYINVK-IRPLENNOQALGTGLRGDEKEDSVDIIEHKRN 376
Db 364 MHGO-YLLTL--RKGGNNKILIKLEHGDGKYGSDPL-----TSSSVELLNNHN 410
0y 269 KYTSMPREP- ----RSRDKVOHNEWYIGESRPAVEAEPMKEKSDSFLVDCST 322
Db 310 -----KKPKPTVANNNGMNNMMSIQNNNEWYMGDSIREVNEK-LRPAIDDTFLVLRDASTK 363

```

RESULT 4

US-08-208-887A-2
Sequence 2, Application US/08208887A
Patent No. 5677421
GENERAL INFORMATION:
APPLICANT: Schlessinger, Joseph
APPLICANT: Skolnick, Edward Y.
APPLICANT: Margolis, Benjamin L.
TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: 10036-2711
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/208,887A
FILING DATE: 11-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-063
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-208-887A-2

Query Match	7.3%;	Score 147;	DB 1;	Length 724;
Best Local Similarity	24.7%;	Pred. No. 2.1e-05;		
Matches	58;	Conservative	35;	Mismatches 73;
			Indels	69;
			Gaps	12

```

QY 176 PEVGMQSIILRDSIELEAEKPNHO-----RKPESTHLENNONGET 220
    ||||| : : : : : |||
Db 211 PEVOSSEYIDL-LKILIRSPSIPHOYWLTYLLKHFKEKLSOTSSKNLNAVISEIF 268
    ||||| : : : : : |||

QY 221 -PLAISSSSFTTNNH-----SVONDRHKGGMQPCSPORCOPCASPHEINTLPY 268
    ||||| : : : : : |||
Db 269 SPMLFRFSASDSMDTENLIKVIETLITSENMERQ-----PALALP 309
    ||||| : : : : : |||

QY 269 KYTSMRPEPK-----RSDRKDVQHNEMWYIGEYSKQAVEAEAKENKDGSEFLVDCSTK 322
    ||||| : : : : : |||
Db 310 -----KPKPPTTVANNGNNNNNSLQNMENWYMGDISISREEVNEK-LDRTADGSEFLVROASTK 363
    ||||| : : : : : |||

QY 323 SKEEPLYALVAYENKVVYWK-IRFLERLQOALGTGLRGDEKPEFSVEDIIEHYKN 376
    ||||| : : : : : |||
Db 364 MHGD-VYTLTL---RGGGNKKLIKLFHRGCKGFSDDL---TFSSVVELLNHYN 410
    ||||| : : : : : |||

```

RESULT 5

```

US-08-539-005-2
Sequence 2, Application US/08539005
Patent No. 585686
GENERAL INFORMATION:
APPLICANT: Schliesinger, Joseph
APPLICANT: Skolnick, Edward Y.
APPLICANT: Margolis, Benjamin L.
TITLE OF INVENTION: NOVEL EXPRESSION CLONING METHOD FOR
TITLE OF INVENTION: IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINASE
TITLE OF INVENTION: KINASES AND NOVEL TARGET PROTEINS
NUMBER OF SEQUENCES: 50
CORRESPONDENCE ADDRESS:
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: 10036-2711
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/539,005
FILING DATE: 4-OCT-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/167,035
FILING DATE: 16-DEC-1993
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-062
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-539-005-2

```

Query Match	7.38;	Score 147;	DB 2;	Length 724;
Best Local Similarity	24.78;	Pred. No. 2.1e-05;		
Matches	58;	Conservative	35;	Mismatches 73;
				Indels 69;
				Gaps 12;

```

0Y 176 PEVQMGMSQISLRDSEVLEAEKVPHNQ-----RKPESTHLENNQGEI 220
      ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 211 PEVQSSSEYIOL--LKKLIRBPSLPHQWYTLQYLKHFHFLSTOSSKNLLNARVSEIF 268
      | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 221 -PLAIISSSFTTSMH-----SVQNDRHGGMOPCSPOBCQPPASCSPHENILPY 268
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 269 SPMLFRFSAASDNTENLIKVIETLSTENMEFO-----PAPALPP----- 309
      | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 269 KITSRRPFEPK-----KSDKRDQVHNWYIGEYSROAVGEAFKKNKDGSEFLYRDCSTK 322
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 310 -----KPKRPPTTVANGNNNNNNNSLONAAWYGDLSREEVNKR-LRDPADGFEFLYRDCSTK 363
      | | | | | | | | | | | | | | | | | | | | | | | | | |
0Y 323 SKEEYVLAVEYENKVYVNVK-IRLEFRNQOFGTLGJGDEKFPDSVEDIIEHYN 376
      | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 364 MHGD-YTILTL--RKGGNNKLKIFHRGKRGKGFSDPL-----TFSSVVELLNHYNN 410
      | | | | | | | | | | | | | | | | | | | | | | | | | |

```

Sequence 5, Application US/09280598
Patent No. 6391584
GENERAL INFORMATION:
APPLICANT: Schlusser, Joseph
APPLICANT: Scholnik, Edward Y.
APPLICANT: Margolis, Benjamin L.
TITLE OF INVENTION: A NOVEL EXPRESSION-CLONING METHOD FOR
IDENTIFYING TARGET PROTEINS FOR EUKARYOTIC TYROSINE
KINASES AND NOVEL TARGET PROTEINS
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/280,598
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/252,820
FILING DATE: 02-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Coruzzi, Laura A.
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7683-067
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 724 amino acids
TYPE: amino acid
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-09-280-598-5

Query Match 7.3%, Score 146.5; DB 4; Length 724;
Best Local Similarity 24.7%, Pred. No. 2.3e-05;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;

QY 176 PEVQGPSQISLRDSEVLEAKVPHNQ-----RKPESTHLLNONTQET- 220
| | | | | : : : : : | | | | |
DB 211 PEVQSSSEYQL-LKKLIRSPSIHQYWLTLQYLKHFPLKSQTSKNLLARVLSLTF 268
| | | | | : : : : : | | | | |
QY 221 -PLAIISSSFTTNSH-----SVQNRDRHGMOPCSPORCOPASPSPHENILPY 268
| : : : : : | : : : : : | | | | |
DB 269 SPMLFRFSAASDNTENLIKVEILISTEMNRO-----PAPALPP----- 309
| : : : : : | : : : : : | | | | |
QY 269 KYTSRRPEPRK-----KSDRDVQHNENYIGEYSRQAVEAFMKENKDGSLVDCSTK 322
| : : : : : | : : : : : | : : : : : | | | | |
DB 310 -----KPKRPTTVANNNGNNNNNSLQNAEMWGDISREYNEK-LRDTADGTFVLVDASTK 363
| : : : : : | : : : : : | : : : : : | | | | |
QY 323 SKEEPYVAVFENKVVYVK-IRFLERNOQFALGTGLRDEKFDSEVEDIIEHYKN 376
| : : : : : | : : : : : | : : : : : | | | | |
DB 364 MHGD-YTILTL---RKGNNKLIKIRHROKYGFSDDL-----TFSSVELINHYRN 410
| : : : : : | : : : : : | : : : : : | | | | |

RESULT 7
US-08-390-874C-12
Sequence 12, Application US/08390874C
GENERAL INFORMATION:
APPLICANT: Klippel, Anke

APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: A Constitutively Active
Phosphatidylinositol 3-kinase and Uses Thereof
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/390,874C
FILING DATE: 17-FEB-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Dow, Karen B.
REGISTRATION NUMBER: 29,684
REFERENCE/DOCKET NUMBER: 02307K-057000US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 722 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-390-874C-12

Query Match 7.2%, Score 144.5; DB 3; Length 722;
Best Local Similarity 21.4%, Pred. No. 3.6e-05;
Matches 75; Conservative 40; Mismatches 95; Indels 141; Gaps 17;

QY 141 PLPPRPPLITLPPKYYQPLPEPERSR-----PLPSQRTPEVQGPSQISLR 188
| | | | | : | | | | | : | | | | |
DB 85 PTPKPRP---PR---PLPVAGSSKTEADTPEQALPLPDLAEQFAPPDV-APLLIKL 136
| | | | | : | | | | | : | | | | |
QY 189 DLSEV-LEAEVPHNQKPESTHLE-----NONTQETPLAIISSSFTTNSHVNOR 239
| | | | | : | | | | | : | | | | |
DB 137 EAIEKKGLCESTLYRTQSSNPAPELRQLDDCAASVDELMIDVHLADAFKRYLADLPN- 195
| | | | | : | | | | | : | | | | |
QY 240 DHRGMOPCSP-----QRCOPASC-----SPHENILPYKY----- 270
| | | | | : | | | | | : | | | | |
DB 196 -----FVIPAAYNEMMSLAQELQSEDFCQLLKLIRLPNPHQCMTLTLQYLKHF 248
| | | | | : | | | | | : | | | | |
QY 271 -----TSW----- 273
| | | | | : | | | | | : | | | | |
DB 249 KLSQASSKNLLNARVLSLTFSPVLFPPAASDNTENLIKVEILISTEMNROAPALP 308
| : : : : : | : : : : : | | | | |
QY 274 -RPPRP-----KSDRDVQHNENYIGEYSRQAVEAFMKENKDGSLVDCSTKSE 326
| : : : : : | : : : : : | : : : : : | | | | |
DB 309 PKRPPTTVANNNGNNNNNSLQDAEMWGDISREYNEK-LRDTADGTFVLVDASTKMG 366
| : : : : : | : : : : : | : : : : : | | | | |
QY 327 PYVLAVFENKVVYVK-IRFLERNOQFALGTGLRDEKFDSEVEDIIEHYKN 376
| : : : : : | : : : : : | : : : : : | | | | |
DB 367 -YTILTL---RKGNNKLIKIRHROKYGFSDDL-----TFNSVELINHYRN 409
| : : : : : | : : : : : | : : : : : | | | | |

RESULT 8
US-09-265-772-12
Sequence 12, Application US/09265772
Patent No. 6300111
GENERAL INFORMATION:
APPLICANT: Klippel, Anke
APPLICANT: Williams, Lewis T.
TITLE OF INVENTION: A Constitutively Active


```

; TITLE OF INVENTION: Phosphatidylinositol 3-Kinase and Uses Thereof
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,772
; FILING DATE: 10-MAR-1999
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/390,874
; FILING DATE: 17-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Dow, Karen B.
; REGISTRATION NUMBER: 29,684
; REFERENCE/DOCKET NUMBER: 02307K-057010US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 722 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-265-772-12

```

```

Query Match      7.2%; Score 144.5; DB 4; Length 722;
Best Local Similarity 21.4%; Pred. No. 3.6e-05;
Matches 75; Conservative 40; Mismatches 95; Indels 141; Gaps 17;

```

```

QY 141 PLPPPPPLTLPPKYPOLPEPESSR-----PPLSGRHFFPEVQGMPSQISLR 188
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 85 PPKPKPP-----PR---PLPVAFGSSKTEADTEGOALPLPLABQFAPDV-APPLIKTL 136
      ||| | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 189 DLSEV--LEAEKYPHNRKPESTHLE-----NNTQEIPLAIISSSTTTNHSYQR 239
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 137 EAEKGLGECSTLYRTQSSSNPAELRQLDDCAASVDLEMDIVHADAKRYLADLPN- 195
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 240 DHRGMQPCSP-----ORCOPASC-----SPHENILPYKY----- 270
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 196 -----PIIPVAVYVEMMSLAQELDSPEDCIOLKKLIRLPNIPRQCHLITLOYLKHF 248
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 271 -----TSW----- 273
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 249 KLSQASSKNILNARVLEIFSPVLFRRPAAASDNTHELIKAEILISTEMENRQAPALP 308
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 274 -RRPF-----KRSRKDYQHNEMWYIGESRQAVEAEAFKEKKDSFLYRDCSTSKKE 326
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 309 PKPKPTTYANNMNMNMSLQDAEWYWGDISREYNEK--LRDTADGTFTLRDASTYKMGD 366
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 327 PYLAIVFENKYYNVK-IRFLERNQOFALGTGLRGDEKFPDVEDITEHYKN 376
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 367 -YTLTL-----RKGNKLIKIFHRDGKYGFSPL-----TFNSVELLNHYRN 409
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 9
US-08-391-615-2
; Sequence 2, Application US/08391615
; Patent No. 5550054
; GENERAL INFORMATION:
; APPLICANT: Witte, Owen
; APPLICANT: Tsukada, Satoshi

```

```

; APPLICANT: Saffran, Douglas
; APPLICANT: Rawlings, David
; TITLE OF INVENTION: HEMATOPOIETIC RESTRICTED TYROSINE KINASE
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FLEHR, HOHNACH, TEST, ALBRITTON & HERBERT
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/391,615
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/006,449
; FILING DATE: 21-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Rowland, Berttram I
; REGISTRATION NUMBER: 20,015
; REFERENCE/DOCKET NUMBER: A-57583-1/BIR UCLA 232-1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 398-3249
; TELEFAX: (415) 398-3249
; TELEEX: 910 277299 FHT UR
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-391-615-2

```

```

Query Match      7.1%; Score 143; DB 1; Length 659;
Best Local Similarity 23.4%; Pred. No. 4.4e-05;
Matches 60; Conservative 27; Mismatches 71; Indels 98; Gaps 11;

```

```

QY 125 RSNINIGDASVRKNKIPPLPPRPLTLPPKYYOLPEPESSRPLSGRHFPEVQGMPSQ 184
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 171 RNSGLKPGSSHRKTRKPLPTPEEDQILK--PLPEPTA-----AP 210
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 185 ISLRDSEVLEA-EKVPHNO-----RRPESTHLENONFOEIPLAIISSSFTTSHSVON 238
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 211 ISTTELKRYVALDYMPNMANDQLRKGEYFLESN--LFW-----WRA 254
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 239 RDHRGMQPCSPQRCOPASCSPHENILPYKYTSWRPPEKRSRDRDYQHNEMWYIGESYR 298
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 255 RDKNG-----QEGYIPSNYIT-----DAEDSIEMEYSHMTR 288
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 299 QAVEEAPMKENKDGSPLYRDCSTKSKKEEYVLAVFENKYYNKKIRPLENNOQFALGTGL 358
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 289 SQAEOLLKQKGGKGGFIVRSKAGK-----YTVSY-----FAKSTG- 325
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 359 RGDEKFDVEDITEHY 374
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 326 -----EPQGVIRHY 334
      || | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

RESULT 10
US-08-426-509A-8
; Sequence 8, Application US/08426509A
; Patent No. 6326469
; GENERAL INFORMATION:
; APPLICANT: Ullrich, Axel
; APPLICANT: Gishitsky, Mikhail
; APPLICANT: Sures, Irman G.

```

```

; TITLE OF INVENTION: NOVEL MEGAKARYOCYTIC PROTEIN
; TITLE OF INVENTION: TYROSINE KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York,
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,509A
; FILING DATE: 21-APR-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/232,545
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-0074-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: No. 6326469e
; US-08-426-509A-8

```

Query Match 6.9%; Score 140; DB 4; Length 659;

Best Local Similarity 22.7%; Pred. No. 8.5e-05; Matches 58; Conservative 29; Mismatches 71; Indels 98; Gaps 11;

```

QY 125 RSONIKGDASYRNKKIPLPPRPPLITLPKKYQPLPPEPSSRPPLSQRHTRPEVQGMPSQ 184
Db 171 RNSGLKGGSSHRKTKKPLPRPEEDQILKK--PLPPEPAA-----Ap 210
QY 185 ISLSDLSEVLEA-EKVPHNQ-----RKPESTHLENONTOEIPLAISSSFTTSNHSYON 238
Db 211 VSTSELKKVVALYDYMPMNANDLQLRKGDDEFILLESN--LPW-----WRA 254
QY 239 RDHNGMGOPCSPORCOPPASCSPHENILPYKYTSWRRPFRKSRDKDYQHNEWYIGESR 298
Db 255 RDKNG-----OEGYIPSNYV-----EAEDSTIEMEWYSKHNTR 288
QY 299 QAVEAFAFKENKKGDFLVRCSTSKSEEPYLAIFYENKYNVVKIRPLERNQOFALGTGL 358
Db 289 SQAEQLLKQEGKEGGFIVRDSKAGK-----YTVSV-----FAKSTG- 325
QY 359 RGDEKFDSEVDIIIEHY 374
Db 326 -----DPOGVIRHY 334

```

RESULT 11
PCT-US95-05008-8
Sequence 8, Application PC/TUS9505008

GENERAL INFORMATION:
APPLICANT: Sugen, Inc.
APPLICANT: 515 Galveston Drive
APPLICANT: Redwood City, California 94063-4720
APPLICANT: United States of America
APPLICANT: Wissenschaften E.V.

```

; APPLICANT: Hofgarten Str. 2
; APPLICANT: Munchen 80539
; APPLICANT: Germany
; TITLE OF INVENTION: Novel Megakaryocytic Protein Tyrosine
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/05008
; FILING DATE: 24-APR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/232,545
; FILING DATE: 22-APR-1994
; TELECOMMUNICATION INFORMATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7683-074
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)790-9090
; TELEFAX: (212)869-9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 659 amino acids
; TYPE: amino acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; PCT-US95-05008-8

```

Query Match 6.9%; Score 140; DB 5; Length 659;

Best Local Similarity 22.7%; Pred. No. 8.5e-05; Matches 58; Conservative 29; Mismatches 71; Indels 98; Gaps 11;

```

QY 125 RSONIKGDASYRNKKIPLPPRPPLITLPKKYQPLPPEPSSRPPLSQRHTRPEVQGMPSQ 184
Db 171 RNSGLKGGSSHRKTKKPLPRPEEDQILKK--PLPPEPAA-----Ap 210
QY 185 ISLSDLSEVLEA-EKVPHNQ-----RKPESTHLENONTOEIPLAISSSFTTSNHSYON 238
Db 211 VSTSELKKVVALYDYMPMNANDLQLRKGDDEFILLESN--LPW-----WRA 254
QY 239 RDHNGMGOPCSPORCOPPASCSPHENILPYKYTSWRRPFRKSRDKDYQHNEWYIGESR 298
Db 255 RDKNG-----OEGYIPSNYV-----EAEDSTIEMEWYSKHNTR 288
QY 299 QAVEAFAFKENKKGDFLVRCSTSKSEEPYLAIFYENKYNVVKIRPLERNQOFALGTGL 358
Db 289 SQAEQLLKQEGKEGGFIVRDSKAGK-----YTVSV-----FAKSTG- 325
QY 359 RGDEKFDSEVDIIIEHY 374
Db 326 -----DPOGVIRHY 334

```

RESULT 12
US-08-391-615-6
Sequence 6, Application US/08391615
Patent No. 5550054
GENERAL INFORMATION:


```

; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 474 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-729-416C-1

```

```

Query Match          6.4%; Score 129.5; DB 3; Length 474;
Best Local Similarity 24.7%; Pred. No. 0.00052;
Matches 72; Conservative 37; Mismatches 112; Indels 71; Gaps 16;

```

```

QY 112 RLRRVVKPIRSDYRSQINIGDASVRK--NKIRLPPRPRLITLTKKQPLRPPESS-----165
DB 201 RMOSLDEPWTEE-----EDGSDHPYNSIPSKMPGGFELDTRLKPRHAPDQAQFAG 254
QY 166 --RPPLSQRH---TFPE-VQGMPSQISLRDLSEVLEAE-KVPHNQKPESTHLENNQTQ 218
DB 255 KEQTYQGRRLGDTFGEDMQQTPRLRGSSDIYSTPEGKLVHAPTEGAP--TYV-----NTQ 308
QY 219 ET-----PLAISSSFTTSMHVSQVNDHRCGM--QPCSP-----ORCPAPASCSPHEN 264
DB 309 QIPQAMPAAVSSAESPRKDLFDMKPFEDALKNQPLGPIVLSKAASVEICISPV-----362
QY 265 ILPYKYSMPRPKPSDRKQVQHNEMVYIGEYSRQAVEEAFMKNKDGSLVYDCSTKSK 324
DB 363 -----PRAPDAKMLEELQAEWTWYQGESMRKEAEGILL---EKDGDVLR---KST 405
QY 325 EEP--YLAVFYENKYVNVKIRFLERNOQFALGTGLRGDEKPDVEDIIEHY 374
DB 406 TNGSFVLGMHNGQAKHLLVDP-----GTIRKDRVFDSISHLINHH 450

```

```

RESULT 15
US-08-729-416C-7
; Sequence 7, Application US/08729416C
; Patent No. 6013767
; GENERAL INFORMATION:
; APPLICANT: NAKAMURA, TAKESHI
; TITLE OF INVENTION: NOVEL BRAIN-SPECIFIC ADAPTER MOLECULE GENE
; TITLE OF INVENTION: THEREOF, AND ANTIBODY THERETO
; NUMBER OF SEQUENCES: 22
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO, L.L.P.
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/729,416C
; FILING DATE: 11-OCT-1996

```

```

; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: PERRY, GLENN J.
; REGISTRATION NUMBER: 28458
; REFERENCE/DOCKET NUMBER: 7898/225948
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 594 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-729-416C-7

```

```

Query Match          6.4%; Score 129.5; DB 3; Length 594;
Best Local Similarity 24.7%; Pred. No. 0.00072;
Matches 72; Conservative 37; Mismatches 112; Indels 71; Gaps 16;

```

```

QY 112 RLRRVVKPIRSDYRSQINIGDASVRK--NKIRLPPRPRLITLTKKQPLRPPESS-----165
DB 321 RMOSLDEPWTEE-----EDGSDHPYNSIPSKMPGGFELDTRLKPRHAPDQAQFAG 374
QY 166 --RPPLSQRH---TFPE-VQGMPSQISLRDLSEVLEAE-KVPHNQKPESTHLENNQTQ 218
DB 375 KEQTYQGRRLGDTFGEDMQQTPRLRGSSDIYSTPEGKLVHAPTEGAP--TYV-----NTQ 428
QY 219 ET-----PLAISSSFTTSMHVSQVNDHRCGM--QPCSP-----ORCPAPASCSPHEN 264
DB 429 QIPQAMPAAVSSAESPRKDLFDMKPFEDALKNQPLGPIVLSKAASVEICISPV-----482
QY 265 ILPYKYSMPRPKPSDRKQVQHNEMVYIGEYSRQAVEEAFMKNKDGSLVYDCSTKSK 324
DB 483 -----PRAPDAKMLEELQAEWTWYQGESMRKEAEGILL---EKDGDVLR---KST 525
QY 325 EEP--YLAVFYENKYVNVKIRFLERNOQFALGTGLRGDEKPDVEDIIEHY 374
DB 526 TNGSFVLGMHNGQAKHLLVDP-----GTIRKDRVFDSISHLINHH 570

```

```

Search completed: April 21, 2003, 12:39:08
Job time : 15.5906 secs

```

GenCore version 5.1.4_P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:36:32 ; Search time 13.4451 Seconds
(without alignments)
2114.581 Million cell updates/sec

Title: US-09-856-061-4
Perfect score: 2020
Sequence: 1 FQNFSLPKRNSWPRINSATG.....GLRGEKEDVEDIEHYKN 376

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 288829 seqs, 75613885 residues
Total number of hits satisfying chosen parameters: 288829

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications -AA: *
1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep: *
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB pep: *
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep: *
4: /cgn2_6/ptodata/2/pubpaa/US06_PUB pep: *
5: /cgn2_6/ptodata/2/pubpaa/US07_PUB pep: *
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep: *
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB pep: *
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep: *
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep: *
10: /cgn2_6/ptodata/2/pubpaa/US05_PUBCOMB pep: *
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep: *
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep: *
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep: *
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2009	99.5	428	9	US-09-966-955A-4
2	2009	99.5	443	9	US-09-966-955A-2
3	1694	83.9	353	9	US-09-966-955A-6
4	238	11.8	46	10	US-09-864-761-38652
5	214	10.6	40	10	US-09-864-761-44023
6	149	7.4	688	9	US-10-081-980B-1
7	147	7.3	724	10	US-09-962-929-2
8	146.5	7.3	724	10	US-09-962-929-4
9	144.5	7.2	945	8	US-08-965-272-2
10	144.5	7.2	945	12	US-10-121-882-2
11	143	7.1	659	9	US-10-045-202-4
12	140	6.9	659	9	US-09-977-260-8
13	140	6.9	659	9	US-10-045-202-2
14	140	6.9	659	9	US-09-977-261-8
15	140	6.9	659	10	US-09-977-269-8
16	126	6.2	597	9	US-10-038-010-22
17	125.5	6.2	533	10	US-09-908-805B-63
18	125	6.2	593	10	US-09-920-021A-3

ALIGNMENTS

20	120	5.9	939	10	US-09-226-248B-29	Sequence 29, Appl
21	120	5.9	939	10	US-09-801-368-380	Sequence 380, App
22	119	5.9	620	9	US-09-977-260-9	Sequence 9, Appl
23	119	5.9	620	9	US-09-977-261-9	Sequence 9, Appl
24	119	5.9	620	10	US-09-977-269-9	Sequence 9, Appl
25	115	5.7	675	9	US-10-186-399-3	Sequence 3, Appl
26	115	5.7	675	9	US-09-977-260-4	Sequence 4, Appl
27	115	5.7	675	9	US-09-977-261-4	Sequence 4, Appl
28	115	5.7	675	10	US-09-977-269-4	Sequence 4, Appl
29	112.5	5.6	552	10	US-09-880-192-50	Sequence 50, Appl
30	111	5.5	542	10	US-09-908-805B-44	Sequence 44, Appl
31	111	5.5	655	10	US-09-205-658-57	Sequence 57, Appl
32	111	5.5	655	10	US-09-844-353A-57	Sequence 57, Appl
33	111	5.5	944	10	US-09-964-238-2	Sequence 2, Appl
34	110	5.4	1341	9	US-10-076-622-565	Sequence 565, App
35	110	5.4	1341	12	US-10-007-805-555	Sequence 555, App
36	110	5.4	1349	9	US-10-076-622-573	Sequence 573, App
37	110	5.4	1349	12	US-10-007-805-573	Sequence 573, App
38	109.5	5.4	4019	9	US-09-854-133-425	Sequence 425, App
39	109.5	5.4	4019	10	US-09-738-973-425	Sequence 425, App
40	107.5	5.3	365	9	US-10-029-180-62	Sequence 62, Appl
41	107.5	5.3	1191	10	US-09-921-099-2	Sequence 2, Appl
42	107.5	5.3	1191	10	US-09-921-099-4	Sequence 4, Appl
43	107.5	5.3	1325	10	US-09-864-761-35612	Sequence 35612, A
44	106	5.2	673	10	US-09-727-384-4	Sequence 4, Appl
45	105.5	5.2	479	10	US-09-881-752A-312	Sequence 312, App

RESULT 1
US-09-966-955A-4
; Sequence 4, Application US/09966955A
; Patent No. US2002015563A1
; GENERAL INFORMATION:
; APPLICANT: Perez-Villar, Juan J.
; APPLICANT: Chang, Han
; APPLICANT: Yang, Wen-Pin
; APPLICANT: Wu, Yuli
; APPLICANT: Whitney, Gene S.
; APPLICANT: Kanter, Steven B.
; TITLE OF INVENTION: Identification and Cloning of a Full-Length Human
; TITLE OF INVENTION: Clk-related Gene, MIST (Mast Cell Immunoreceptor)
; TITLE OF INVENTION: Signal Transducer)
; FILE REFERENCE: 3053-4113US1
; CURRENT APPLICATION NUMBER: US/09/966,955A
; CURRENT FILING DATE: 2001-09-28
; PRIOR APPLICATION NUMBER: 60/237030
; PRIOR FILING DATE: 2000-09-29
; NUMBER OF SEQ ID NOS: 52
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 428
; TYPE: PRT
; ORGANISM: HUMAN
; FEATURE:
; OTHER INFORMATION: HUMAN MIST SPICE VARIANT CLONE #7, AMINO ACID
; OTHER INFORMATION: SEQUENCE
US-09-966-955A-4

Query Match 99.5% Score 2009; DB 9; Length 428;
Best Local Similarity 99.5% Pred. No. 8.7e-142;
Matches 374; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 FQNFSLPKRNSWPRINSATGQYORMNKPLDWERNEFAVLGAKGSHDDYDDPELAMEE 60
Db 19 FQNFSLPKRNSWPRINSATGQYORMNKPLDWERNEFAVLGAKGSHDDYDDPELAMEE 78
QY 61 TWOSIKLPPRPKESYADTHYFKVAMDPLPLDRTSISISIQPTWNTOTRLERVDKPI 120
Db 79 TWOSIKLPPRPKESYADTHYFKVAMDPLPLDRTSISISIQPTWNTOTRLERVDKPI 138


```
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 38652
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005599.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 5.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.9
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 5.4
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 5.7
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 5.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.4
; OTHER INFORMATION: SWISSPROT HIT: Q60787, EVALUATE 9.00e-10
US-09-864-761-38652

Query Match      11.8%; Score 238; DB 10; Length 46;
Best Local Similarity 100.0%; Pred. No. 2.5e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 311 DGSFLVDCSTKREPYVLAFYENKYNKIRFLERNOOPALGT 356
Db 1 DGSFLVDCSTKREPYVLAFYENKYNKIRFLERNOOPALGT 46
```

```
RESULT 5
US-09-864-761-44023
; Sequence 44023, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; FILE REFERENCE: Aecomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44023
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005599.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.72
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.88
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.68
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.47
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.87
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 0.81
; OTHER INFORMATION: SWISSPROT HIT: P25502, EVALUATE 3.70e+00
US-09-864-761-44023

Query Match      10.6%; Score 214; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 34 RNFAVLGAGKSHDDDDYDDPELRMEETWOSIKILPARI 73
DB 1 RNFAVLGAGKSHDDDDYDDPELRMEETWOSIKILPARI 40

RESULT 6

US-10-081-980B-1

; Sequence 1, Application US/10081980B
; Publication No. US20030041337A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Gibbs, E. Michael
; APPLICANT: McNeish, John D.
; TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form Of
; FILE REFERENCE: PC10206AGR
; CURRENT APPLICATION NUMBER: US/10/081,980B
; PRIOR FILING DATE: 2002-08-23
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 1
; LENGTH: 688
; TYPE: PRT
; ORGANISM: mus musculus
US-10-081-980B-1

Query Match 7.4%; Score 149; DB 9; Length 688;
Best Local Similarity 21.7%; Pred. No. 0.0035;
Matches 76; Conservative 39; Mismatches 96; Indels 140; Gaps 17;
QY 141 PLPPRPLTLTKKYQPLPEPSSR-----PPLSORHTFPVQGMPSQISLR 188
DB 85 PPKPRP-----PR---PLPVAGSSKTEADTEQALPLPDLAEQFAPPV-APPLIKILL 136
QY 189 DLSEV--LEAEKVPNQRPESTHLE-----NONTQEIPLAIISSSFTTNSHVQR 239
DB 137 EAIEKKGECSTLYRTQSSNPALRQLDCAASVDELMIDVHLADAFKRYLADLPN- 195
QY 240 DHRGGMQPCSP-----QRCOPASC-----SPENILPKYK----- 270
DB 196 -----PVIPAVYVNMMSLQDELQSPEDCIQLKKLRLRNIPHCWLTQYLLKHPF 248
QY 271 -----TSM----- 273
DB 249 KLSQASSKMLNARVLSIFSPVLFPPAASSDNTHELIKAEILISTEMNERQAPALP 308
QY 274 -RPPRP-----KSDRKDVQHNWYIGESRQAVEAEAFMKENKDGSLVRCSTKSKEE 326
DB 309 PPKPRPTTVANNMNNMNSLQDAEWYMGDISREVEK-LRDTADGTFLVRDASTKMGD 367
QY 327 PYLAVFYENKYYNVK-IRFLERNOQFALGTGLRGDEKFDSEVIEDIEHKKN 376
DB 368 -YTLTL---RKGNKMLKIFHRDGKYGSDPL---TNSVVELINHYRN 410

RESULT 7

US-10-081-980B-3

; Sequence 3, Application US/10081980B
; Publication No. US20030041337A1
; GENERAL INFORMATION:
; APPLICANT: Pfizer Inc.
; APPLICANT: Gibbs, E. Michael
; APPLICANT: McNeish, John D.
; TITLE OF INVENTION: Transgenic Animals Containing A Dominant Negative Mutant Form Of
; FILE REFERENCE: PC10206AGR
; CURRENT APPLICATION NUMBER: US/10/081,980B
; PRIOR FILING DATE: 2002-08-23
; NUMBER OF SEQ ID NOS: 12
US-10-081-980B-3

QY 141 PLPPRPLTLTKKYQPLPEPSSR-----PPLSORHTFPVQGMPSQISLR 188
DB 85 PPKPRP-----PR---PLPVAGSSKTEADTEQALPLPDLAEQFAPPV-APPLIKILL 136
QY 189 DLSEV--LEAEKVPNQRPESTHLE-----NONTQEIPLAIISSSFTTNSHVQR 239
DB 137 EAIEKKGECSTLYRTQSSNPALRQLDCAASVDELMIDVHLADAFKRYLADLPN- 195
QY 240 DHRGGMQPCSP-----QRCOPASC-----SPENILPKYK----- 270

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 3
; LENGTH: 724
; TYPE: PRT
; ORGANISM: mus musculus
US-10-081-980B-3

Query Match 7.4%; Score 149; DB 9; Length 724;
Best Local Similarity 21.7%; Pred. No. 0.0035;
Matches 76; Conservative 39; Mismatches 96; Indels 140; Gaps 17;

QY 141 PLPPRPLTLTKKYQPLPEPSSR-----PPLSORHTFPVQGMPSQISLR 188
DB 85 PPKPRP-----PR---PLPVAGSSKTEADTEQALPLPDLAEQFAPPV-APPLIKILL 136
QY 189 DLSEV--LEAEKVPNQRPESTHLE-----NONTQEIPLAIISSSFTTNSHVQR 239
DB 137 EAIEKKGECSTLYRTQSSNPALRQLDCAASVDELMIDVHLADAFKRYLADLPN- 195
QY 240 DHRGGMQPCSP-----QRCOPASC-----SPENILPKYK----- 270
DB 196 -----PVIPAVYVNMMSLQDELQSPEDCIQLKKLRLRNIPHCWLTQYLLKHPF 248
QY 271 -----TSM----- 273
DB 249 KLSQASSKMLNARVLSIFSPVLFPPAASSDNTHELIKAEILISTEMNERQAPALP 308
QY 274 -RPPRP-----KSDRKDVQHNWYIGESRQAVEAEAFMKENKDGSLVRCSTKSKEE 326
DB 309 PPKPRPTTVANNMNNMNSLQDAEWYMGDISREVEK-LRDTADGTFLVRDASTKMGD 367
QY 327 PYLAVFYENKYYNVK-IRFLERNOQFALGTGLRGDEKFDSEVIEDIEHKKN 376
DB 368 -YTLTL---RKGNKMLKIFHRDGKYGSDPL---TNSVVELINHYRN 410

RESULT 8

US-09-962-929-2

; Sequence 2, Application US/09962929
; Patent No. US20020115056A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Pinn S.
; APPLICANT: Soerensen, Annette B.
; APPLICANT: Nielsen, Anne A.
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Diseases Associated wit
; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,929
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/668,644
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO: 2
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-962-929-2

Query Match 7.3%; Score 147; DB 10; Length 724;
Best Local Similarity 21.7%; Pred. No. 0.005;
Matches 76; Conservative 38; Mismatches 97; Indels 140; Gaps 17;
QY 141 PLPPRPLTLTKKYQPLPEPSSR-----PPLSORHTFPVQGMPSQISLR 188
DB 85 PPKPRP-----PR---PLPVAGSSKTEADTEQALPLPDLAEQFAPPV-APPLIKILL 136
QY 189 DLSEV--LEAEKVPNQRPESTHLE-----NONTQEIPLAIISSSFTTNSHVQR 239
DB 137 EAIEKKGECSTLYRTQSSNPALRQLDCAASVDELMIDVHLADAFKRYLADLPN- 195
QY 240 DHRGGMQPCSP-----QRCOPASC-----SPENILPKYK----- 270

Db 196 -----PVIPVAVNEMMSLAOELSPEDCIOLKLRPLPNIHQWTLQYLKHF 248
QY 271 -----TSW----- 273
Db 249 KLSQASSKULNARVSEIFSPVLFPPAASNTHELIKAIELLISTENNEROPAPALP 308
QY 274 -RPPFP-----KSDRKDVQHNEMYIGEXSROAVEAFKRNKDGSLVRDCSTKSEE 326
Db 309 KPPKRTTYANNMNMNMNSLODAEWYWGDISREVENK -LRDTADGTFVLVRDSTKMHGD 367
QY 327 PYLAVFYENKYNVK-IRFLERNOQFALGTGRGDEKFDPSVEDIIEHYKN 376
Db 368 -YTLT---PRKGNMNLKIKIFHRDGKYGFSDDL---TFNSVVELINHYRN 410

RESULT 9
US-09-962-929-4
; Sequence 4, Application US/09962929
; Patent No. US20020115058A1
; GENERAL INFORMATION:
; APPLICANT: Pedersen, Finn S.
; APPLICANT: Soerensen, Annette B.
; APPLICANT: Nielsen, Anne A.
; TITLE OF INVENTION: Methods for Diagnosis and Treatment of Diseases Associated with A
; FILE REFERENCE: A-70004/RMS/DCF
; CURRENT APPLICATION NUMBER: US/09/962,929
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: US 09/668,644
; SOFTWARE: Patent version 3.1
; SEQ ID NO 4
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-962-929-4

Query Match
Best local Similarity 7.3%; Score 146.5; DB 10; Length 724;
Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;

QY 176 PEVOGMPQISLRDLSEVLEAEKVPANO-----RKPESTHLENOQTQEI- 220
Db 211 PEVOSSSEYIQL--IKLIRSPISPHQYMLTQYLKHFKLSTQTSKLNARVSEIF 268
QY 221 -PLAIISSSTFTSNH-----SVQNRDRHGGMQPCSPQRCOPPASCSEHENTLPY 268
Db 269 SPMLFRFSAASSDNTENLIVIEILISTENNERO-----PAPALP----- 309
QY 269 KYTSWRPPFPK-----RSDRKDVQHNEMYIGEXSROAVEAFKRNKDGSLVRDCSTK 322
Db 310 -----KPPKRTTYANNMNMNMNSLODAEWYWGDISREVENK -LRDTADGTFVLVRDSTK 363
QY 323 KEEPYLAVFYENKYNVK-IRFLERNOQFALGTGRGDEKFDPSVEDIIEHYKN 376
Db 364 MHGD-YTLT---RKGNMNLKIKIFHRDGKYGFSDDL---TFNSVVELINHYRN 410

RESULT 10
US-08-965-272-2
; Sequence 2, Application US/08965272
; Patent No. US20020019514A1
; GENERAL INFORMATION:
; APPLICANT: Glimcher, Laurie H. et al.
; TITLE OF INVENTION: CITRA-INTERACTING PROTEINS AND METHODS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts

COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/965,272
FILING DATE:
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kara, Catherine J.
REGISTRATION NUMBER: 41,106
REFERENCE/DOCKET NUMBER: HUI-029
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 945 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-965-272-2

Query Match
Best local Similarity 7.2%; Score 144.5; DB 8; Length 945;
Matches 64; Conservative 22; Mismatches 80; Indels 81; Gaps 12;

QY 53 DPELMEETWQSIK--ILPAPRIKESYADTHYFKVAMDTPLD---TTSISIGOPTW 107
Db 5 DESIRKEEQOQHAGVAPQPLKEP-----FASLOSPPFDTPATTAVAVAATTT 56
QY 108 NFOIRLEVDKPSRDVRSQNIKGDASVRKNKIP---LPPRPPLITLPKKYOPLPPEPES 164
Db 57 TTTTITTT-----TAQDEKKRPPALPPPPPLAKPPPSQOPPPP-- 97
QY 165 SREPLSGRHTPEVOGMPQIS-LRDLSEVLEAEK-----VPHNORP- 206
Db 98 --PP-----PSPASLKLASVLEGGKYCRGTGAAVSTRPGPLPTTQYSPG 142
QY 207 -----ESTHLENOQTQEIPLAISSTFTSNHVSQNRDRHGGMQ---PCSPQRCQ 254
Db 143 PPSGATLPTSAAPSAGSPPSASSSOSTSG-GPMARERRAGEEYPVGPWTPTOPP 201
QY 255 PPASCSP 261
Db 202 PPLSLPP 208

RESULT 11
US-10-121-882-2
; Sequence 2, Application US/10121882
; Patent No. US20020146806A1
; GENERAL INFORMATION:
; APPLICANT: Glimcher, Laurie H. et al.
; TITLE OF INVENTION: CITRA-INTERACTING PROTEINS AND METHODS
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```

RESULT 12
US-10-045-202-4
: Sequence 4, Application US/10045202
: Publication No. US20030040461A1
: GENERAL INFORMATION:
: APPLICANT: Bristol-Myers Squibb
: TITLE OF INVENTION: MODULATORS OF BRUTON'S TYROSINE KINASE
: TITLE OF INVENTION: INTERMEDIATES AND METHODS FOR THEIR IDENTIFICATION AND USE IN THE TREATMENT OF B-CELL LYMPHOMAS
: TITLE OF INVENTION: PREVENTION OF OSTEOPOROSIS AND RELATED DISEASE STATES
: FILE REFERENCE: D0032 NP
: CURRENT APPLICATION NUMBER: US/10/045, 202
: CURRENT FILING DATE: 2002-06-27
: PRIOR APPLICATION NUMBER: 60/242,471
: PRIOR FILING DATE: 2000-10-23
: NUMBER OF SEQ ID NOS: 12
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 4
: LENGTH: 659
: TYPE: PRT
: ORGANISM: Murine
: US-10-045-202-4

```

RESULT 14

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:33:12 ; Search time 14.3724 Seconds
(without alignments)
2515.001 Million cell updates/sec

Title: US-09-856-061-4

Perfect score: 2020

Sequence: 1 FGNFSLPKRNSWPRINSATG.....GLRGDEKFDVEDIEHYKN 376

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332.5	16.5	533	2	A56110
2	315.5	15.6	533	2	B56110
3	149	7.4	724	2	A38747
4	146.5	7.3	724	2	A38748
5	146	7.2	724	2	A38749
6	143	7.1	659	2	I49553
7	140	6.9	659	2	I49444
8	140	6.9	659	2	A45184
9	138.5	6.9	981	1	FOMVGM
10	136.5	6.8	841	2	A43254
11	131.5	6.5	485	2	T37550
12	131	6.5	723	2	B38749
13	127	6.3	663	1	TVWVRR
14	125.5	6.2	728	2	H59435
15	125	6.2	377	2	S08636
16	125	6.2	585	2	A46209
17	125	6.2	593	1	UN0805
18	125	6.2	595	1	A53551
19	125	6.2	597	1	A53593
20	123.5	6.1	1196	2	T23832
21	123	6.1	461	2	A57463
22	122	6.0	822	1	TYVHFE
23	121.5	6.0	1893	1	A40262
24	121	6.0	1171	2	T12956
25	120.5	6.0	893	2	G88551
26	120	5.9	593	2	JC5167
27	120	5.9	939	2	S28394
28	120	5.9	996	2	T47518
29	119.5	5.9	2649	2	T51023

30	119	5.9	480	2	JC7552
31	119	5.9	620	1	S33253
32	118.5	5.9	1392	2	T51947
33	118	5.8	546	2	S52314
34	117.5	5.8	289	2	T52354
35	117.5	5.8	667	2	T27672
36	116.5	5.8	1162	2	T49191
37	116	5.7	545	2	S52313
38	115	5.7	675	2	S60612
39	115	5.7	796	2	E96654
40	115	5.7	1290	2	A36466
41	114.5	5.7	625	1	A43030
42	114.5	5.7	630	1	T01380
43	114.5	5.7	741	2	T40095
44	114	5.6	517	2	A43807
45	114	5.6	1128	2	A43960

ALIGNMENTS

RESULT 1

A56110
tyrosine phosphoprotein SLP-76 - human
C:Species: Homo sapiens (man)
C:Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
C:Accession: A56110
R:Jackman, J.K.; Motto, D.G.; Sun, Q.; Tanemoto, M.; Turk, C.W.; Peltz, G.A.; Koretz, J. Biol. Chem. 270, 7029-7032, 1995
A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with
A:Reference number: A56110; MUID:95221245; PMID:7706237
A:Accession: A56110
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-533 <OAC>
A:Cross-References: GB:U20158; MID:9806765; PIDN:AAC50135.1; PID:9806766
C:Genetics:
A:Gene: GDB:ICP2; SLP-76
A:Cross-References: GDB:1230199; OMIM:601603
A:Map position: 5q33.1-5qter
C:Superfamily: SH2 homology
C:Keywords: phosphoprotein
F:422-520/Domain: SH2 homology <SH2>

Query Match	Score	332.5	DB 2	Length	533
Best Local Similarity	25.3%				
Matches	113	Conservative	49	Mismatches	128
Indels	157	Gaps	16		
OY	20	GGYQKRNKPLDWRNFAVLGAKGSHDDYDDPELRMEETWQSTIKILPAPRIKSEYA	79		
DB	126	GDESPN-----EEEAAYED-----DADYPPPSNDEALQN-SILPAKFPNS---	169		
OY	80	DTHYFVAMDPLDTRTSISIGQPTWNTQRLERVDKPISRDVSNQIKGDASVRKKK	139		
DB	170	NSMY-----KRYOLRP-----EPSSR-----IDRPPSGKTRQGP-----	186		
OY	140	IPLPRLPLTLTP-----KKYOLRP-----EPSSR-----	166		
DB	187	PVPFQPMALPPLPPAGRNHSPPLPPQTNHPEFSRHHKTAKLAPASIDNSTKPLDLR	245		
OY	167	-----PPLSQRTFP-----EVOGMPQSILDLSEVLEAEKV	199		
DB	246	SLAPFREPFTLGKRPFDKPSIPAGRSGLHLPKIQPLPPTTERHRSRSLPGKKR	305		
OY	200	P-----HNRKPESTHLEKNT-----QETPLAISSSTT	230		
DB	306	PVPKHWGPDREKREDDVHQRPLPQALPLPSSNTFSPSRSTKPSPMNPLPSSHMGARS	365		
OY	231	TSNHSVONDRHGMQPCSPQ--RCOPPASCSPHENTILPYKTTSMWPPPKRSDDKRDVQH	288		
DB	366	ENSSSF---PQASLSLPPYSQGSNRPPIRACGRNPLPLP-NKPRPSP--AEEENSIN	419		
OY	289	NEWYIGYSRQAVAEAFMKENKDGSEFLVRDCSTKSKKEEYVLAVEYNKYVNVKIRFLER	348		

Db 420 EEWVSYITPREAEALARKINQDGFILVRDSSKKTTPYLVMLVKQKYNIGIRYCKE 479
 QY 349 NQFALGTGLGDEKEDFSYEDIIIEHYK 375
 Db 480 SQVYLGTGLGKEDFLSVSDIIDYFR 506

RESULT 2

B56110
 tyrosine phosphoprotein SLP-76 - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 28-Apr-1995 #sequence_revision 28-Apr-1995 #text_change 05-Nov-1999
 C:Accession: B56110
 R:Jackson, J.K.; Motto, D.G.; Sun, Q.; Tanemoto, M.; Turck, C.W.; Peltz, G.A.; Koretzky, J. Biol. Chem. 270, 7029-7032, 1995
 A:Title: Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with C
 A:Reference number: A56110; MUID:95221345; PMID:7706237
 A:Accession: B56110
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-533 <IAC>
 A:Cross-references: GB:020159; NID:9806767; PIDN:AAC52189.1; PID:9806768
 C:Superfamily: SH2 homology
 C:Keywords: phosphoprotein
 F:422-520/Domain: SH2 homology <SH2>

Query Match 15.6%; Score 315.5; DB 2; Length 533;
 Best Local Similarity 30.0%; Pred. No. 1.5e-14;
 Matches 121; Conservative 44; Mismatches 169; Indels 69; Gaps 18;

QY 20 GQYRMNPKPLDMERNFAVLGAGSHDDYDDPELMEETWOSIKILPARPI--KKESE 77
 Db 126 GDYESPN-----EEQALVDDAA-----DYEPSPNNEALQS-SILPNSFNTISM 172
 QY 78 YAD-THYFKVAMDPL-PLDRTSISIGPTWNTQRLERVDKPISRDVSQNIKGDAV 135
 Db 173 YIDRPTGKVSQPPVPLRPKPALP-PLPTGRNHSPLSPHPNHEERSGNNK----- 226
 QY 136 RKNKIPPL-----PPRLTLTKKYPPL-----PPPESSRPLSRHNFPEVQ-GM 181
 Db 227 -TALPLASIDRSTKPLPLDRSLAPDRREPFLGKKPPSPKSAPLGRH-LPKIQKPL 284
 QY 182 PSQSLRLDSEVL-----EAEKVPNORKPESTHLE---NONTQELPLAISSSFTTSN 233
 Db 285 PPADNRHNRERLGPVYTRKRSVPRHGRGPRREDEDVQORPLPQSLSLSMSSNTSPS 344
 QY 234 HSYQ-----NRDRGMQPCS--PQRCQPPASCSP--HENTLPKYYS 272
 Db 345 RSVQSSKNTFPPLAHMPCAFESNIGFQSSASLPPYFGCGNRPPLRSEGRNLPPLVPN 404
 QY 273 WRPPPRKSDKDVQHNWYIGESRQAVEAFKKNKDSFLVRDCTSKKEEYVLAV 332
 Db 405 -RPPPSGEEETPLDEWYVSYITPREAEALARKINQDGFILVRDSSKKTTPYLVMLV 463
 QY 333 FYENKVVNKRIFLERNOQFALGTGLGDEKEDFSYEDIIIEHYK 375
 Db 464 LYKDKVNIQIRQESQVYLGTGLGKEDFLSVSDIIDYFR 506

RESULT 3

A38747
 phosphatidylinositol 3-kinase (EC 2.7.1.-) 85k chain - mouse
 C:Species: Mus musculus (house mouse)
 C>Date: 03-Aug-1992 #sequence_revision 03-Aug-1992 #text_change 16-Jul-1999
 C:Accession: A38747
 R:Escobedo, J.A.; Navarokasattusas, S.; Kavanaugh, W.M.; Milfay, D.; Fried, V.A.; William Cell 65, 75-82, 1991
 A:Title: cDNA cloning of a novel 85 kD protein that has SH2 domains and regulates binding
 A:Reference number: A38747; MUID:91191564; PMID:1849460
 A:Accession: A38747
 A:Molecule type: mRNA
 A:Residues: 1-724 <ESC>

A:Cross-references: GB:M60651
 C:Comment: This protein binds a phosphotyrosine-containing sequence of ligand-activat
 phosphatidylinositol at position 3 of the inositol ring.

C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:333-428/Domain: SH2 homology <SH2A>
 F:624-718/Domain: SH2 homology <SH2>

Query Match 7.4%; Score 149; DB 2; Length 724;
 Best Local Similarity 21.7%; Pred. No. 0.01;
 Matches 76; Conservative 39; Mismatches 96; Indels 140; Gaps 17;

QY 141 PLPPRPPLITPKKYQPLPEPPESR-----PLSQHRTPEVQGMPSQSLR 188
 Db 85 PTPKPRP-----PR---PLPAPGSSKTEADTEQALPLPDLAEQFAPRDY-APPLILKL 136
 QY 189 DLSEY--LEAKVPNORKPESTHLE-----NONTQELPLAISSSFTTSNHSVQR 239
 Db 137 EAIEKKGLECSLYRTQSSNPDELRLDQDAASVDLEMDIVLADAFKRYLADLPN- 195
 QY 240 DHRGMQPCSP-----ORCQPPASC-----SPHENILPKYK----- 270
 Db 196 -----PVITVAVYNNEMSLAQELQSPEDCTQLKLILRLPNIPHOQWLTLQILKHF 248
 QY 271 -----TSW----- 273
 Db 249 KLSQASSKNILNARYLSIFSPVLEFRPAASDNTHEILKAILISTENMERQAPALP 308
 QY 274 -RPPPP-----KSDRKDVQHNWYIGESRQAVEAFKKNKDSFLVRDCTSKKEE 326
 Db 309 PKPKPTTVANNMNNMSLQDAEWYMGDISREEVNEK-LRDTADGFLVHDASTKMHGD 367
 QY 327 PYVLAVFENKVVNKR-IRFLERNOQFALGTGLGDEKEDFSYEDIIIEHYK 376
 Db 368 -YTLTL--RRGNNKLIKIFHRDQKYGFSDDL-----TENSVELINHYRN 410

RESULT 4

A38748
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85k chain - human
 C:Species: Homo sapiens (man)
 C>Date: 24-Jan-1992 #sequence_revision 24-Jan-1992 #text_change 16-Jul-1999
 C:Accession: A38748; S28402
 R:Skolnik, E.Y.; Margolis, B.; Mohammadi, M.; Lowenstein, E.; Fischer, R.; Drepps, A. Cell 65, 83-90, 1991
 A:Title: Cloning of P13 kinase-associated p85 utilizing a novel method for expression
 A:Reference number: A38748; MUID:91191565; PMID:1849461
 A:Accession: A38748
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-724 <SKO>
 A:Cross-references: GB:M61906
 R:Panayotou, G.; Bax, B.; Gout, I.; Federwisch, M.; Wroblewski, B.; Dhand, R.; Fry, M. EMBO J. 11, 4261-4272, 1992
 A:Title: Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2 domain
 A:Reference number: S28402; MUID:9049176; PMID:1330535
 A:Accession: S28402
 A:Status: preliminary
 A:Molecule type: protein
 A:Molecule type: protein
 C:Genetics:
 A:Gene: GDB:PIK3R1
 A:Cross-references: GDB:127604; OMIM:171833
 A:Map position: 5q12-5q13
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:333-428/Domain: SH2 homology <SH2A>
 F:624-718/Domain: SH2 homology <SH2>

Query Match 7.3%; Score 146.5; DB 2; Length 724;
 Best Local Similarity 24.7%; Pred. No. 0.015;
 Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;

```

QY 176 PEVOGMQSI1RDIJSEVLEAEKVPBHQ-----RKRESHILLNEMTQEI 220
    |||| :||: :||: :||: :||: :||: :||: :||: :||: :||:
Db 211 PEVOSSSEYIOL--LKKILIRSPISLPHQYWTLLQYLIKHFELLSQISSKNLLNARLSEIF 268

QY 221 -PLAIISSSSFTTNSH-----SVONDRHRCGMOPCSPORCOPASCSPHENILPY 268
    |||| :||: :||: :||: :||: :||: :||: :||: :||:
Db 269 SPMLFRFSAASNDNTENLIKVIETLISJEMNERQ-----PAPALP----- 309

QY 269 KYTSNRPFPR-----RSDRKDVQHNEMWTIGEYSKQAVEAEAFMKENKQSEFLVDCSTK 322
    |||| :||: :||: :||: :||: :||: :||: :||: :||:
Db 310 -----KPKRPFTTVANGNNNNMNSLONAEWGTGDISREEVNER-LRDTADGTGFLVADASTK 363

QY 323 SKEPPYLAAYENKVVYVWK--IRFLERNQOFALGTGLGDEKFDSEVEDIIEHYH 376
    |||| :||: :||: :||: :||: :||: :||: :||: :||:
Db 364 MHGD-YILLTL--RKGGNNKLKIFHRDGKTFGSDPL-----TFSSVLEIINHYN 410

```

RESULT 5
A38749
3-phosphatidylinositol kinase (EC 2.7.1.-) 85K chain alpha - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 14-Feb-1992 #sequence_revision 14, Feb-1992 #text_change 05-Nov-1999
C:Accession: A38749
R:Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, A
Cell 65, 91-104, 1991
A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine k
A:Reference number: A38749; MUID:91191567; PMID:1707345
A:Accession: A38749
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1724 <ONS>
A:Cross-references: GB:M61746; GB:M61745; NID:g163476; PIDN:AAA79511.1; PID:g163477
C:Superfamily: SH2 homology
C:Keywords: phosphotransferase
F:335-428/Domain: SH2 homology <SH2A>
F:624-718/Domain: SH2 homology <SH2>

Query Match	7.2%;	Score 146;	DB 2;	Length 724;
Best Local Similarity	22.4%;	Pred. No. 0.016;		
Matches 77; Conservative	49;	Mismatches 103;	Indels 114;	Gaps 16

```

Oy      136  RKNKPLPPPPPLTLTPKKYOLPPPESSR-----PLSOLHFPV-----178
Db      80  KKSPTPKPPPPPLPVPAPSPKSTLEADSEQASTLDLAQRPVPVAPPLIKLVEAI 139

Oy      179  -QGMPQISLR-----DLSEVLEAEKVP-----HNOKPESTHLENOQT 220
Db      140  EKKGECSSTLYKROSSNPBELRLDLDCTASLDLEMDVHLVDAFKRRLDLPN-PVI 198

Oy      221  PLAISSSFT-----TSNSHV 236
Db      199  PVAVSELSILAPVQSSSEYIQLKKLIRSPTRHQYWLTLQYLKHFKPLQSTSSKNL 258

Oy      237  QNRDHGGMQPCSPQPCPASCSPH-ENLP-----YKYSWR-----PPPKR-- 280
Db      259  LNA--RVLSLSPLEFFFPAASENEHLKIIEILISTENRQDAPALPKPKPKPT 316

Oy      281  -----SDRKVQHNEHVGEGYSQAVAEAEAKENKSGFLVPRCSTPKSEEPVLAIFY 334
Db      317  VANNGMNNMNSLDQAEWWDGISNEBVEVAK-LRPTAGCTFLVYRASRKMGMG-TYTL-- 372

Oy      335  EKKYVNVK-IRPLENQOALGTGLRGDKEDSDVEDIIEHKH 376
Db      373  -RKGGNNKLIRHEDKGYGFSDP---TNSVVELIHNHN 410

```

protein-tyrosine kinase (EC 2.7.1.112) emb - mouse
N:Alternate names: B-cell progenitor cytoplasmic tyrosine kinase; Bruton agammaglobulinemia
C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Feb-2000

C:Accession: 195553; B45184; JN0471; PC4366
R:Sidners, P.; Muller, S.; Shiels, H.; Jin, H.; Khan, W.N.; Nilsson, L.; Parkinson, I
Ulos, K.G.
J: Immunol. 153, 5607-5617, 1994
A:Title: Genomic organization of mouse and human Bruton's agammaglobulinemia tyrosin
A:Reference number: 149553; MUID:95081608; PMID:7989760
A:Accession: 149553
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-659 <SID>
A:Cross-references: GB:I29788; NID:g625143; PTDN:AA66943.1; PID:g625144
R:Tsukada, S.; Saffran, D.C.; Rawlings, D.J.; Parolini, O.; Allen, R.C.; Kilsak, I.;
Cell 72, 279-290, 1993
A:Title: Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-link
A:Reference number: A45184; MUID:93145329; PMID:8425221
A:Accession: B45184
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-122,'P',124-659 <R5U>
A:Cross-references: GB:I08967; NID:g192233; PIDN:AA37316.1; PID:g192234
A:Experimental source: 702/3 pre-B cell line
A:Note: sequence extracted from NCBI backbone (NCBI:12383) and corrected to corres
R:Yanada, N.; Kawakami, Y.; Kimura, H.; Fukamachi, H.; Baier, G.; Altman, A.; Kato,
Biochem. Biophys. Res. Commun. 192, 231-240, 1993
A:Title: Structure and expression of novel protein-tyrosine kinases, Emb and Emc, in
A:Reference number: JN0471; MUID:93236578; PMID:8476425
A:Accession: JN0471
A:Status: nucleic acid sequence not shown

A:Molecule type: mRNA
A:Residues: 1-66,'A','68-197','W1',198-449,451-659 <YAM>
A:Cross-references: GB:LI0627; NID:g193016
A:Note: this translation is not annotated in GenBank entry MUSEMBX, release 116.0
R:Kojima, T.; Fukuda, M.; Watanabe, Y.; Hamazato, F.; Mkoshiba, K.
Biochem. Biophys. Res. Commun. 236, 333-339, 1997
A:Title: Characterization of the pleckstrin homology domain of Btk as an inositol phosphate
A:Reference number: PC4366; MUID:97382431; PMID:9240435
A:Accession: PC4366
A:Molecule type: mRNA
A:Residues: 1-165 <KOJ>
C:Comment: This protein specifically recognizes the inositol 1,3,4,5-pentakisphosphate molecule to the membrane.

A:Gene: emb;Btk
C:Superfamily: protein-tyrosine kinase tec; pleckstrin repeat homology; protein kinase
D:Keywords: ATP; phosphoprotein; phosphotransferase; tyrosine-specific protein kinase
E:1-131/Domain: pleckstrin repeat homology <PLK>
F:221-269/Domain: SH3 homology <SH3>
F:281-377/Domain: SH2 homology <SH2>
F:400-658/Domain: protein kinase homology <KIN>
F:408-416/Region: protein kinase ATP-binding motif
F:551/Binding site: phosphate (Tyr) (covalent) #status predicted

[illegible]

Db 326 -----EPGVIRHY 334

RESULT 7

149444

SH3 binding protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999

C:Accession: 149444

R:Ref: R. Mayer, B.J. Cicchetti, P. Baltimore, D.

Sentence 259, 1157-1161, 1993

A:Title: Identification of a ten-amino acid proline-rich SH3 binding site.

A:Reference number: 149444; MUID:93174278; PMID:8438166

A:Accession: 149444

A:Molecule type: mRNA

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-559 <RES>

A:Cross-references: GB:L14543; NID:9293267; PIDN:AAA37121.1; PID:9293268

A:Superfamily: pleckstrin repeat homology

F:25-128/Domain: pleckstrin repeat homology <PLK>

F:201-210/Region: proline-rich SH3 binding

Query Match

Best Local Similarity 6.9%; Score 140; DB 2; Length 559;

Matches 93; Conservative 54; Mismatches 156; Indels 168; Gaps 18;

56 LMEETWOSIKILPAPRKESSEYADTHYKVM-----TPPLD 95

82 MRAEETSNVFPFKITHSKHRTWFSASSEDEKRSMAFVREIGHEKKEPLD 141

96 TRTSISIQPTWNTQRLERVDKPISRDVRSQINIKDASVKKNKIP--LPFRP----- 147

142 TSDSS-----DTDSFGAVERPIDISLSTYPMNEDYEHEDDDSYLEPSPMKLE 195

148 -LITPKKYQPLPEPSSRPPLS---QRHTF-----PEVQMSQIL-RDL 190

196 DALTPPAPYPP-PPVPPVRKPAFSDLPRAHSTSKSPSPPLPPPPKRGLPDGTGAPEDA 254

191 SEVLEAEKVPNHQKREESTHLENONTQEIPLAIS----- 225

255 KDALGLRVEQLRVPAIRPRMSDPPMSNVPLVRKHPCRDSVNGLEPMTWGHGTS 314

226 --SSFTTSHSVQNRD-----HRGKQP-----CSPQRCQ--- 254

315 SVSSTTAAVATSRNCDLKSFLHSSRGPTSEPPVPANKPKFLKIAEPPSPRAKFA 374

255 --PRASCSPHENILPYKXTSMRPPRPKRSDKRDVH----- 288

375 PVPVAPRPPVQKMPKMPKPAVPAVLPPEPTPLPRLQSRPPDGSFSGFSFEKARQPSQ 434

289 -----NEWYIGYSRQAVEAFMK-----ENKGSFLVDCSTKSKE 325

435 ADTGEDEDEDEYKVPRLNSVFNTESECEVERLFKATDPGCEPDGLXCIKRNSTKSGK 494

326 EPPYLAAYE--NKYVANKIRFLERNOQFALGTGLRDEKDSVEDITHY 374

495 --VLVWDESSNKNRYRI--FEKDSKFY----LEGEVLFAVSQSWVEHY 536

RESULT 8

A45184

protein-tyrosine kinase (EC 2.7.1.112), nonreceptor type, BTK - human

N:Alternate names: alk; B cell progenitor kinase; Bruton agammaglobulinemia tyrosine kin

C:Species: Homo sapiens (man)

C:Date: 30-Apr-1993 #sequence_revision 02-Jul-1996 #text_change 21-Jul-2000

C:Accession: 137212; 154541; 139457; S28912; A45184

R:Ref: Y. Haire, R.N. Litman, R.T. Fu, S.M. Nelson, R.P. Kratz, J. Kornfeld, S.J.

Proc. Natl. Acad. Sci. U.S.A. 91, 9067-9066, 1994

A:Title: Genomic organization and structure of Bruton agammaglobulinemia tyrosine kinase

A:Reference number: 137212; MUID:94377492; PMID:8090769

A:Accession: 137212

A:Status: nucleic acid sequence not shown; translation not shown; not compared with c

A:Molecule type: DNA

A:Residues: 1-659 <OHNT>

A:Cross-references: EMBL:U10087; NID:9126232; PIDN:AAB60639.1; PID:9517438

A>Note: Only Intron-exon junctions are shown

R:Ref: R. Rohrer, J. Parolino, O. Belmont, J.W. Conley, M.E.

Immunogenetics 40, 319-324, 1994

A:Title: The genomic structure of human BTK, the defective gene in X-linked agammaglo

A:Reference number: 154541; MUID:95012452; PMID:7927535

A:Accession: 154541

A:Status: nucleic acid sequence not shown; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-659 <ROH>

A:Cross-references: GB:L1572; NID:9632958; PIDN:AAA61479.1; PID:9632960

R:Ref: R. Hagmann, T.L. Chen, Y. Rosen, F.S. Kwan, S.P.

Hum. Mol. Genet. 3, 1743-1749, 1994

A:Title: Genomic organization of the Btk gene and exon scanning for mutations in pat

A:Reference number: 139457; MUID:95152493; PMID:7880320

A:Accession: 139457

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-278, 281-659 <HAG>

A:Cross-references: EMBL:U13433; NID:9575888; PIDN:AAC51347.1; PID:9575890

R:Ref: R. Vetterle, D. Vorechovsky, I. Sideras, P. Holland, J. Davies, A. Fliinter, F. Ham

Nature 361, 226-233, 1993

A:Title: The gene involved in X-linked agammaglobulinemia is a member of the src fam

A:Reference number: 528912; MUID:93140868; PMID:8380905

A:Accession: 528912

A:Molecule type: mRNA

A:Residues: 1-659 <VEN>

A:Cross-references: GB:X58957; NID:9312466; PIDN:CAA41728.1; PID:9312467; GB:U78027;

R:Ref: R. Tsukada, S. Saffran, D.C. Rawlings, D.J. Parolinski, O. Allen, R.C. Klisak, I.

Cell 72, 279-290, 1993

A:Title: Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-link

A:Reference number: A45184; MUID:93145329; PMID:8425221

A:Accession: A45184

A:Status: not compared with conceptual translation

A:Molecule type: DNA

A:Residues: 1-19, 'R', 21-190, 'A', 199-206, 'T', 208-210, 'I', 212-218, 'E', 220-238, 'E', 240-2

A:Cross-references: GB:U78027; GB:L35265; NID:92281904

A:Experimental source: erythroleukemia cell line K562

A>Note: sequence extracted from NCBI backbone (NCBI:123835)

A:Genetics:

A:Gene: GDB:BTK; AGMX1; IMD1

A:Cross-references: GDB:120542; OMIM:300300

A:Map position: Xq21.33-Xq22

A:Introns: 47/3; 80/3; 103/3; 131/1; 174/1; 196/3; 259/2; 280/2; 298/3; 325/2; 368/1;

C:Superfamily: protein-tyrosine kinase; tyrosine-specific protein kinase

C:Keywords: ATP; phosphotransferase; pleckstrin repeat homology; protein kinase

F:2-131/Domain: pleckstrin repeat homology <PLK>

F:221-269/Domain: SH3 homology <SH3>

F:281-377/Domain: SH2 homology <SH2>

F:408-416/Region: protein kinase ATP-binding motif

Query Match

Best Local Similarity 6.9%; Score 140; DB 2; Length 659;

Matches 58; Conservative 29; Mismatches 71; Indels 98; Gaps 11;

125 RSQNTKGDASVKKNKIPPLPPRLITLPKKTQPLPEPSSRRPLSQHNTPEVQMPQ 184

171 RNSGLKPGSSHRKTKRPPPEEDQILK--PLPEPA-----Ap 210

185 ISLRDLSEVLA-EKVPNQ-----RKPESTHLENONTQEIPLAISSSFTTSHSVON 238

211 VSTSELRKVVALLVDMPPNANDQLRKDEFFILESN--LPW-----wRA 254

239 RDHKGMPQSPQRCQPPASCPHENILPYKYTSMRPPRPKRSDKRDVQHNEWYIGESR 298

255 RDKNG-----QEGYIPSNVYV-----EAEDSIEMEWESKIMTR 288

299 QAVEAEFVKENKDGSLVRCSTSKSEPPYLAAYFENKVVNVAIRFLERNOQFALGTGL 358

Db 289 SOAEOLLEKQEGGFIYVDSKAGK-----YTVSV-----FAKSTG- 325
 QY 359 RGDEKFDSDIEDIIEHY 374
 Db 326 -----DPQGVIRHY 334

RESULT 9

FOMVGM

gag-abl polyprotein - Abelson murine leukemia virus

N:Contains: amino end of core shell protein p30; core protein p15; inner coat protein p1

C:Species: Abelson murine leukemia virus

A:Note: host Mus sp. (mouse)

C:Date: 14-Nov-1983 #sequence_revision 09-Sep-1994 #text_change 11-Jun-1999

C:Accession: A03931; A00627; A93955

R:Reddy, E.P.; Smith, M.J.; Strinivasan, A.

Proc. Natl. Acad. Sci. U.S.A. 80, 3623-3627, 1983

A:Title: Nucleotide sequence of Abelson murine leukemia virus genome: structural similar

A:Reference number: A93955; MUID:83221648; PMID:6304726

A:Accession: A03931

A:Molecule type: DNA

A:Residues: 1-981 <RED>

A:Cross-references: GB:J02009; NID:9331887; PIDN:AAA46471.1; PID:9331888

A:Note: the authors translated the codon GGA for residue 186 as Glu

R:Reddy, E.P.; Smith, M.J.; Strinivasan, A.

Proc. Natl. Acad. Sci. U.S.A. 80, 7372, 1983

A:Reference number: A93980

A:Contents: annotation; erratum, residues 588-746

C:Genetics:

A:Gene: gag-abl

C:Superfamily: Abelson murine leukemia virus gag-abl polyprotein; protein kinase homolog

C:Keywords: ATP; core protein; oncogene; phosphotransferase; polyprotein; transforming F

F:1-131/Product: core protein p15 #status predicted <p15>

F:132-215/Product: inner coat protein p12 #status predicted <p12>

F:216-235/Region: amino end of core shell protein p30

F:248-338/Domain: SH2 homology <SH2>

F:361-621/Domain: protein kinase homology <KIN>

F:369-377/Region: protein kinase ATP-binding motif

F:392/Active site: Lys #status predicted

Query Match

Best Local Similarity 6.9%; Score 138.5; DB 1; Length 981;

Matches 62; Conservative 40; Mismatches 90; Indels 79; Gaps 16;

QY 143 PPP--RPLTLPPKYYOPLP-----EPSSRPPLSQRH--TFPEVQMPQSISLRD-- 189

Db 93 PPPWVPEVH-PKPPPLPPSAPSLPEPLPSTPPRSSLYPALTPSLGAKPKQVLSDSG 151

QY 190 --LSEVLEAKYPHNQRKESTHLENONTQELPLAISSSFTTSHSVQNRDHRGC-MQ 246

Db 152 GPLLDLTEDPPPRDPPRPPS-----DRDNGNGEAT 183

QY 247 PC-----SPQ-----RCQPPASCSPHENILPYK-----YTSWRPPEPKRSDR----- 283

Db 184 PGEAPDEPSPMARSRLRGREPPVADSTSQAFPLRTGGNGQLQYW--PF--SSSDLYITTV 240

QY 284 KVVQHNWITIGEYSROAVEAFPMKKNKDGSLVRDCTSKSEEPYLAFTENKYYNVAI 343

Db 241 NSLEKSHWYHGFSVNRMAA-EYLLSSGINSFLVRE--SSSPGQGISLIRGEGRVYHYRI 297

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

Db 298 NTASDGKLY-----VSSSESRFNTLAELVHHH 323

QY 344 RFLERNQGFALGTGLGDEKFDSDIEDIIEHY 374

A:Title: corkscrew encodes a putative protein tyrosine phosphatase that functions to

A:Reference number: A43254; MUID:92346711; PMID:1638629

A:Accession: A43254

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-841 <PER>

A:Cross-references: GB:M94730; NID:9157144; PID:9157145

A:Experimental source: embryo

A:Note: sequence extracted from NCBI backbone (NCBIN:109964, NCBIPI:109965)

C:Genetics:

A:Gene: FlyBase:csW

A:Cross-references: FlyBase:FBgn0000382

C:Superfamily: protein-tyrosine-phosphatase, nonreceptor type 6; protein-tyrosine-pho

C:Keywords: phosphoprotein; phosphoric monoester hydrolase; tyrosine-specific phospho

F:6-101/Domain: SH2 homology <SH2>

F:111-203/Domain: SH2 homology <SH2>

F:252-634/Domain: protein-tyrosine-phosphatase homology <PRP>

F:583/Active site: Cys (phosphocysteine intermediate) #status predicted

F:589/Binding site: substrate phosphate (Arg) #status predicted

Query Match

Best Local Similarity 6.8%; Score 136.5; DB 2; Length 841;

Matches 54; Conservative 39; Mismatches 67; Indels 71; Gaps 9;

QY 170 SQRTPEVQGMPSQISLRLDSEVLEAEKVPNNQ-----RK 206

Db 3 SRWFHPTISG-----IEAKLDEQGFDSFLARLSSNGCAFTLSVRRN 49

QY 207 ESTHLENONTQELPLAISSSFTTSHSVQNRDHRGKQPCSPQ--RCQPPASCSPHEN 264

Db 50 EYTH--IKIONNDFEDLYGGEKFAFLPELVQYMMENGELERNGQAIETKQPLICA---- 104

QY 265 ILPYKTSWRPPEPKSRDKDVQHNWYIGEYSFOAVEAFPMKKNKDGSLVRDCTSK 324

Db 105 -----EPTTER-----WFGNLSGKAEKLLIERKNGSFLVRE--SQSK 142

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

QY 325 EEPYLAVEYENKYYNVAIRFLERNOQFALGTGLRDEKFDSDIEDIIEHY 375

Db 143 PDQFVLVSRVTRDQVTHVMIRW--QDKKYDVG---GGESEFGLSELIDHYK 187

Db 260 NEIITKDTPLKDKYSKAPALLPQKPKVS---KQIYOQVSVFSTGKIKESQSLNLTIDT 316
 QY 218 Q-ELPPLAIISSSFSTTNSHVONRHRGMPQSPQRCQPPASCPHENILPYKTTSM--- 273
 Db 317 DIETPLKGS-----ELVSEDFKPPVDPVKIQOI-----LHKONKIIIEKMWISQIRI 364
 QY 274 -----RPPFKRSRDKVDONHENVYIGESRQAVEAEMKENKDGSLVRCOST 321
 Db 365 SKNIEVQRLLDQERHALETIAKKIENNRFILGKRRKA-REAIQKIDN-----LKDLSV 418
 QY 322 KSKEEPPVLAVFENKYNVAKIRLENNQOPALCTGLGDEKFD 365
 Db 419 ---QELFIPESERELKYYELK-----RKDEKLD 443

RESULT 12

B38749
 3-phosphatidylinositol kinase (EC 2.7.1.-) 85K chain B - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 16-Jul-1999
 C:Accession: B38749
 R:Otsu, M.; Hiles, I.; Gout, I.; Fry, M.J.; Ruiz-Larrea, F.; Panayotou, G.; Thompson, A.
 Cell 65, 91-104, 1991
 A:Title: Characterization of two 85 kd proteins that associate with receptor tyrosine k
 A:Reference number: A38749; MUID:91191567; PMID:1707345
 A:Accession: B38749
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-723 <OTS>
 A:Cross-references: GB:M61745; GB:M61746
 C:Superfamily: SH2 homology
 C:Keywords: phosphotransferase
 F:325-420/Domain: SH2 homology <SH2A>
 F:617-706/Domain: SH2 homology <SH2>

Query Match 6.5%; Score 131; DB 2; Length 723;
 Best Local Similarity 21.5%; Pred. No. 0.18;
 Matches 81; Conservative 50; Mismatches 118; Indels 128; Gaps 18;

QY 101 STG-QPTWNTQRLERDQKPSRDVRSQNIKGASVRRKNTLPRLPPRLTLTPKKYOLP 159
 Db 53 SVGMWPGINERT-R-ORGFPGT---YVEFLGPVALAR--PGRPRGPRPLPARPRDGP 104
 QY 160 PEPPSRPPLSQRTFFPVGMPQSPQISLRDSEVLEAKVPHNQRKPESTLLNQNQOE 219
 Db 105 PEGGLTLPDLPEQSPDPV-APPLVVL-----VEATERTGLDSYRPPRAVRTDMSLD 158
 QY 220 IP-----LAISSSFTTS-----NHSYON 238
 Db 159 VEQMDAALSDGVKGFLLAPLVTPEAAEAHRALEAGVGPALERTPLPHALTL 218
 QY 239 R---DHGGMQPCSPQ-----RCPPASCP----- 261
 Db 219 RFLQHLGRVAGRAPPAVALGATFGPRLLRAPPPSPPPGARDGTEPTDFPAL 278
 QY 262 -----HENLIPYKYTSMRPPPKRS-----DRKDYONENYIGESYQA 300
 Db 279 VEKLLQHELEQEAAPAL---PPKPKTIPAPTLGLANGSPSLQDAEMTWGDISKEE 334
 QY 301 VEAFAEMKENKDGSLVRDCKTSKEEPPVLAVFENKYNV-KIRLENNQOPALGTGLR 359
 Db 335 VNEK-LRPTDPTGLVRNASSKIGE-YTLTL---RKGNKKLKVHRDGHYGSSEPL- 388
 QY 360 GDEKFDVYEDILEHKN 376
 Db 389 ---TFCSVVDLITVHRH 402

RESULT 13

TVMYR
 protein-tyrosine kinase (EC 2.7.1.112) fgr - feline sarcoma virus (strain Gardner-Rashed

C:Species: feline sarcoma virus
 A:Note: host fells sp. (cat)
 C>Date: 27-Nov-1985 #sequence_revision 26-May-1995 #text_change 31-Mar-2000
 C:Accession: A00653; A03937
 R:Naharro, G.; Robbins, K.C.; Reddy, E.P.
 Science 223, 63-66, 1984
 A:Title: Gene product of v-fgr onc: hybrid protein containing a portion of actin and
 A:Reference number: A00653; MUID:84097512; PMID:6318314
 A:Accession: A00653
 A:Molecule type: DNA
 A:Residues: 1-663 <NAH>
 A:Cross-references: GB:X00255; GB:X01487; NID:961542; PIDN:CA025063.1; PID:961543
 A:Note: the authors translated the codon GAT for residue 14 as Glu
 C:Comment: This protein is synthesized as a gag-fgr polypeptide.
 C:Genetics:
 A:Gene: fgr
 C:Superfamily: feline sarcoma virus protein-tyrosine kinase fgr: protein kinase homol
 C:Keywords: ATP; autophosphorylation; oncogene; phosphoprotein; phosphotransferase; p
 F:1-118/Region: gag polypeptide similarity
 F:141-268/Region: actin similarity
 F:285-382/Domain: SH2 homology <SH2>
 F:402-660/Domain: protein kinase homology <KIN>
 F:410-418/Region: protein kinase ATP-binding motif
 F:432/Active site: Lys #status predicted

Query Match 6.3%; Score 127; DB 1; Length 663;
 Best Local Similarity 19.1%; Pred. No. 0.31;
 Matches 59; Conservative 44; Mismatches 88; Indels 118; Gaps 12;

QY 142 LPPRPPLTLTPKKYOLP-----PEPPSRPPLSQ----- 171
 Db 101 LPPRPPLTLTPKKYOLP-----PEPPSRPPLSQ----- 171
 QY 172 -----RHPPVQGM-----SQISLRDSEVLE----- 195
 Db 161 FAGDAPPAVFPSTVGRHGVAVGQKDSYGDENQKRGTLTKYPLEHGIVTNMD 220
 QY 196 -AEKVPNQRKPESTHLENO---NTOEIPLAISSSFTTNSHVONRHRGMPQSPQ 251
 Db 221 DMEKIWH-----HFFYNELRVAPREHPVLLLEA-----PLNPK 253
 QY 252 RCPPASCPHENILPYKYTSMRPPPKRSRDKVONHENVYIGESRQAVEAEMKE-NK 310
 Db 254 -----ANREKMTQINFETFNIPSNVAVPDSIQAEVWFGKIGRKDEKRLSPGNA 305
 QY 311 DGSFLVRDCKTSKEEPPVLAV-----FYENKYNVAKIRLENNQOPALGTGLGDEKFD 365
 Db 306 RGAPLVRESST--TKGAYSLSRDMDAERGDHVKHKIRKIDG-----GYITTRAQFN 358
 QY 366 SVEDIDIEHY 374
 Db 359 SVQELVQHY 367

RESULT 14

H59435
 phosphoinositide-3-kinase regulatory beta chain [imported] - huam
 C:Species: Homo sapiens (man)
 C>Date: 03-Jun-2002 #sequence_revision 03-Jun-2002 #text_change 03-Jun-2002
 C:Accession: H59435; A59436
 R:Volinia S.; Patrascu P.; Otsu M.; Hiles I.; Gout I.; Calzolari E.; Bernardi F.; Rooke L.
 Oncogene 7, 789-793, 1992
 A:Title: Chromosomal localization of human p85 alpha, a subunit of phosphatidylinosit
 A:Reference number: H59435
 A:Accession: H59435
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-728 <VOL>
 A:Cross-references: GB:NP_005018; PID:94826908; PIDN:NP_005018.1
 R:Janssen, J.W.; Schleithoff, L.; Bartlam, C.R.; Schulz, A.S.
 Oncogene 16, 1767-1772, 1998
 A:Title: An oncogenic fusion product of the phosphatidylinositol 3-kinase p85beta sub
 A:Reference number: A59436; MUID:98241181; PMID:9582025

A:Accession: A59436
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-728 <JAN>
A:Cross-references: GB:NP_005018; PID:g4826908; PIDN:NP_005018.1

Query Match
Best Local Similarity 20.9%; Score 125.5; DB 2; Length 728;
Matches 72; Conservative 31; Mismatches 101; Indels 141; Gaps 14;

QY 141 PLP-----PPRLITLPKRYQPLPE--PESSRPL-----SQHRTPEV 178
DB 95 PLPARPDGAPPEGLTLP---DLPEQSPDVAAPLVKLVLEAIERTGLDSESHYRPEL 150
QY 179 QGNPQSISLRLSEVLEAEKVPNHQRKESTHLENQTOEIPLAISSFTSNHNSVN 238
DB 151 PAVRTDWSLSDVDQWDTAALA-----DGKSFLLALPAPLVTPESASAEAR 195
QY 239 RDHRGGMQPCSPQ----- 251
DB 196 RALREAGAPVPALEPPTLPALRALTLRFLLQHLGRVARRAPALGPAVRALGATGFL 255
QY 252 RCGPPASCSP-----HENILPYKTSWRPPPKRS-- 281
DB 256 RAPPSPSPPGGAPDGESESPDPFALLVKKLLQEHLEQEVAPPAL---PPKPPKAKP 311
QY 282 -----DRKDVQNEWYIGEYSROAVEAFMKENKDGFLVRCOSTKSKPEPYLAV 332
DB 312 APVYLANGSPPSLQDAEWYWDGISREVEK-LRTPDGTFLVRDASSKIQGE-YTLTL 369
QY 333 FYENKYYNWK-IRLELRNOQFALGTGLRGDEKDFSVEDIIEHYKN 376
DB 370 ---RKGNKMLIKVFHRDGHYGFSEPL---TFCSVVDLINHYRH 407

RESULT 15

S08636
nck protein - human
N:Alternate names: src-related protein
C:Species: Homo sapiens (man)
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 05-Nov-1999
C:Accession: S08636
R:Lehmann, J.M.; Rietmuller, G.; Johnson, J.P.
Nucleic Acids Res. 18, 1048, 1990
A:Title: Nck, a melanoma cDNA encoding a cytoplasmic protein consisting of the src homol
A:Reference number: S08636; MUID:90192089; PMID:2107526
A:Accession: S08636
A:Molecule type: mRNA
A:Residues: 1-377 <LEH>
A:Cross-references: EMBL:X17576; NID:g35014; PID:CAA35599.1; PID:g35015
C:Superfamily: SH3 homology; SH2 homology
F:9-56/Domain: SH3 homology <SH31>
F:113-160/Domain: SH3 homology <SH32>
F:197-247/Domain: SH3 homology <SH33>
F:282-371/Domain: SH2 homology <SH2>

Query Match
Best Local Similarity 23.6%; Score 125; DB 2; Length 377;
Matches 33; Conservative 23; Mismatches 52; Indels 32; Gaps 4;

QY 236 VQNRDRGGMQPCSPQRCPPASCSPHENILPYKYTSWRPPPKRSRDKRDVQHNWYIGE 295
DB 250 MNNPLTSLGLESPPQ-----CDYIRPLTGTGKFGAG-----NPMYYGK 286
QY 296 YSRQAVEEAFMKENKDGSLVRCDSFKSKPEPYLAVFYENKYYNWKIRLELRNOQFALG 355
DB 287 VTRHQAEMLNKGHGDPLIRDSSESPDVSLEKAGKNKHKVQL-----KETVYCI- 341
QY 356 TGLRGDEKDFSVEDIIEHYK 375
DB 342 ----GQRKFSMEIVEHYK 357

Search completed: April 21, 2003, 12:38:38
Job time : 18.3724 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 21, 2003, 12:28:57 ; Search time 8.34525 seconds
(without alignments)
1868.739 Million cell updates/sec

Title: US-09-856-061-4
Perfect score: 2020

Sequence: 1 FQNFSLPKNRSMPRINSATG.....GLRGEKFDVEDIEIHKYK 376

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	332.5	16.5	533	1	LCP2_HUMAN
2	315.5	15.6	533	1	LCP2_MOUSE
3	149	7.4	724	1	P85A_MOUSE
4	146.5	7.3	724	1	P85A_HUMAN
5	146	7.2	724	1	P85A_BOVIN
6	145	7.2	724	1	P85A_RAT
7	143	7.1	659	1	BRK_MOUSE
8	140	6.9	559	1	BRP2_MOUSE
9	140	6.9	659	1	BRK_HUMAN
10	136.5	6.8	845	1	CSW_DROME
11	136	6.7	561	1	3BP2_HUMAN
12	133.5	6.6	724	1	P85B_BOVIN
13	128	6.3	461	1	P55G_HUMAN
14	127	6.3	721	1	CRK_DROME
15	127	6.3	722	1	P85B_MOUSE
16	126.5	6.3	533	1	SOC6_MOUSE
17	126	6.2	461	1	P55G_BOVIN
18	125.5	6.2	728	1	P85B_HUMAN
19	125	6.2	377	1	NCK1_HUMAN
20	125	6.2	585	1	PNB_MOUSE
21	125	6.2	593	1	PNB_HUMAN
22	125	6.2	593	1	PNB_RAT
23	124.5	6.2	217	1	GRAP_HUMAN
24	123.5	6.1	1196	1	ABL1_CAEBL
25	123	6.1	461	1	P55G_MOUSE
26	122	6.0	722	1	P85B_RAT
27	122	6.0	822	1	PER_HUMAN
28	120.5	6.0	893	1	YM92_CAEBL
29	120.5	6.0	847	1	CRK7_HUMAN
30	120	5.9	1490	1	VAV3_HUMAN
31	120	5.9	939	1	SH2_YEAST
32	119	5.9	620	1	ITK_HUMAN
33	119	5.9	1872	1	T2D1_HUMAN

ALIGNMENTS

34	116	5.7	517	1	PCR_MOUSE
35	115	5.7	675	1	BMX_HUMAN
36	115	5.7	1290	1	PIG1_HUMAN
37	114.5	5.7	625	1	ITK_MOUSE
38	114.5	5.7	630	1	TEC_MOUSE
39	114	5.6	1128	1	BEM3_YEAST
40	114	5.6	1190	1	ZO2_HUMAN
41	114	5.6	2440	1	NCR1_HUMAN
42	112.5	5.6	303	1	CRKL_HUMAN
43	112.5	5.6	878	1	VAV2_HUMAN
44	112	5.5	757	1	HT16_HYDAT
45	112	5.5	1290	1	PIG1_RAT

RESULT 1

ID LCP2_HUMAN STANDARD: PRT: 533 AA.

AC Q13094; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 15-JUN-2002 (Rel. 41, Last annotation update)

GN Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP76).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;

RA Jackman J.K., Motto D.G., Sun Q., Tanemoto M., Turck C.W., Peltz G.A., Koretzky G.A., Findeil P.R.; "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein associated with Grb2 in T cells."

RL J. Biol. Chem. 270:7029-7032(1995).

RN (2) SEQUENCE FROM N.A.

RP TISSUE=Prostate;

RA Strausberg R.;

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.

CC -1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND Fyb.

CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES. HIGHLY EXPRESSED ALSO IN T CELL AND MONOCYTIC CELL LINES, EXPRESSED AT LOWER LEVEL IN B CELL LINES.

CC NOT DETECTED IN FIBROBLAST OR NEUROBLASTOMA CELL LINES.

CC -1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION BY ZAP-70.

CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC -1- DATABASE: NAME=PROT; NOTE=PROT 1:1-5(2000);

CC WWW="http://www.ncbi.nlm.nih.gov/prov/guide/1118450040.g.htm"

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.ebi.ac.uk/announcements> or send an email to license@ebi.ac.uk).

DR EMBL; U20158; AAC50135.1; -

DR HSSP; BC016618; AAH16618.1; -

DR HSSP; P12931; 1SHD.

DR Genew; HGNC:6529; LCP2.

DR MIM; 601603; -

DR InterPro; IPR001660; SAM.

DR InterPro; IPR000980; SH2.

DR Pfam; PF00017; SH2; 1.

```

DR ProDom: PD000093; SH2: 1.
DR SMART: SM00454; SAM: 1.
DR SMART: SM00252; SH2: 1.
DR PROSITE: PS50001; SH2: 1.
DR SH2 domain: Phosphorylation.
FT DOMAIN 133 136 POLY-GLU.
FT DOMAIN 198 201 POLY-PRO.
FT DOMAIN 422 530 SH2.
SO SEQUENCE 533 AA; 60188 MW; C5D22F31D36200C8 CRC64;

Query Match 15.5%; Score 332.5; DB 1; Length 533;
Best Local Similarity 25.3%; Pred. No. 1.3e-15;
Matches 113; Conservative 49; Mismatches 128; Indels 157; Gaps 16;

OY 20 GQYQMKPKPLDWERNFPAVLGDAGKSHDDYDDELRLMEETWQSIKILPAPRIESEYA 79
DB 126 GQYSEPN-----EEEEAEVED-----DADYEPPSNDEEALQN-SILPAPFPENS-- 169
OY 80 DTHYKVMADPRLPLDTRTSISIGQPTWQYRLERLVKPKISRDVRSQNIKDAVKNKK 139
DB 170 NSMY-----IDRPPSGKTRPQP----- 186
OY 140 IPLPPRLPLTP-----KKYQPLPP-----EPESRS----- 166
DB 187 -FVPPQRRMAALPPRPAGGNHSPRLPPQTNHEPFRSRNKKAKLPAISIDSTKPLDR 245
OY 167 -----PPLSRHTFP-----EYQMPISOISLRLDSEVLEAEKV 199
DB 246 SLAPDREPFLLGKKRPFSPDKPSIAGKSLGHLKIQKPLPPTTERHERSSPLGKK 305
OY 200 P-----HNRKPESTLLENQW-----QHPLAISSSST 230
DB 306 PPKKGWGPDRREDDVDHQRLPALLPNSSTFPSSRTKPSPMNPPLPSHMGAFS 365
OY 231 TSNHVSQNDHGHGQPCSPQ--RCQPASCSPEHNILPYKTSWMPPEPKRSRDKVOH 288
DB 366 EENSFS---PQASLPPTYSGPSNRNRPRIAGGRNPLPLP-NKRRPSP--AEENSLN 419
OY 289 NEWYIGEXSRQAEAEAFMKENKDGSEFLVDCSTKSKKEEYVLAVFYENKVNVIKIFLER 348
DB 420 EEMVYSYIRPFAEALAKRKINQDGFVLVDSKKTITNYVLAVLKDKVYNIQIRYQKE 479
OY 349 NQQFALGTLRGDEKFDSEYEDIIENKK 375
DB 480 SQVYLLGTGLRGKEDFLSVSDIIDYFR 506

RESULT 2
ICP2_MOUSE STANDARD: PRT; 533 AA.
AC 060787;
DT 16-OCT-2001 (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last annotation update)
DE Lymphocyte cytosolic protein 2 (SH2 domain-containing leucocyte
  protein of 76 kDa) (SLP-76 tyrosine phosphoprotein) (SLP/6).
GN ICP2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sclirognathi; Muridae; Mus.
OX NCBI_TaxID=10990;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell lymphoma;
RX MEDLINE=9521345; PubMed=7706237;
RA Jackson J.K., Mott D.G., Sun Q., Tanemoto M., Turck C.W., Peltz G.A.,
  Koretzky G.A., Finkel P.R.;
RT "Molecular cloning of SLP-76, a 76-kDa tyrosine phosphoprotein
  associated with Grb2 in T cells."
RL J. Biol. Chem. 270:7029-7037(1995).
CC 1- FUNCTION: INVOLVED IN T CELL ANTIGEN RECEPTOR MEDIATED SIGNALING.
  1- SUBUNIT: INTERACTS WITH THE ADAPTER PROTEINS GRB2 AND FTYB.
  1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC

```

CC	-1- TISSUESPECIFICITY: HIGHLY EXPRESSED IN SPLEEN, THYMUS, AND PERIPHERAL BLOOD LEUKOCYTES.
CC	-1- PTM: PHOSPHORYLATED AFTER T-CELL RECEPTOR ACTIVATION (BY SIMILARITY).
CC	-1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.ebi.ac.uk/announcements or send an email to license@ebi.ac.uk).
CC	or send an email to license@ebi.ac.uk).
CC	-----
DR	EMBL; U20159; AAC52189.1; -.
DR	HSSP; P12931; 1SHD.
DR	MCD; MGI:1321402; Lcp2.
DR	InterPro: IPR001660; SAM.
DR	InterPro: IPR000980; SH2.
DR	Pfam; PF00017; SH2; 1.
DR	ProDom; PD000093; SH2; 1.
DR	SMART; SM00454; SAM; 1.
DR	SMART; SM00252; SH2; 1.
DR	PROSITE; PS00001; SH2; 1.
KW	SH2 domain; Phosphorylation.
FT	DOMAIN 422 530
FT	SEQUENCE 533 AA: 60228 MW; 60DB7782C2E60C3 CRC64;
Query Match	15.6%; Score 315.5; DB 1; Length 533;
Best Local Similarity	30.0%; Pred. No. 1.9e-14;
Matches 121:	Conservative 44; Mismatches 169; Indels 69; Gaps 18;
QY	20 GQYGRNKPFLDWERNFAVLGDAGKSHDDYDDPELMEETWOSIKILPARPI--KSE 77
DB	126 GDYESPN-----EEEOALVDDA-----DYEPSPNNEEALOS-SILPNSFHTNSM 172
QY	78 YAD-THYEVANDTFL-PLDRTSTISIQPTWNTQTLERVDKPTISROYRGNIKGDASV 135
DB	173 YIDRPPTGKVSQAPVPLPRKPAFL-PLPTGRNHSPLSPPHNEEPSRSNNK----- 226
QY	136 RKNKPIPLP-----PPRLILPLPKYQPL-----PPEESSRPLSORHTPEVQ--GM 181
DB	227 -TAKLPAPSIDSTPRPLDRSLAPLDREFILGKKRPPSDKRSALGRNH-LPKIQKPL 284
QY	182 PSQISLRDSEVL-----EAEKVPNHQRKESTHLE--NONTQEIPLAISSSFTTSN 233
DB	285 PRAMDREHNEHRLGPTVTRKPSVPRHGRDPRDRENDQDVRPLPQSLPSSMSTPEPS 344
QY	234 HSYQ-----NNDHKGMPGCS--PQRCPAPSCSP--HEMLIPYKTS 272
DB	345 RSQVSSKNTFPLAHMPGAFSSSNTIGFOOSALPPLFFSGGPNRPPLNSEGNNLPLVFN 404
QY	273 WRPPPKRSRDRDOVQNEAYIGEYSQVAEAFMKENKGSFLVDRDCSTKSKKEEYVLAV 332
DB	405 -RPQPSPEEETPLDEEYVSYITRPEEALAKRINDQGTFLVADSSKKTANNPYVLAV 463
QY	333 FYENKYVNVKIRFLERNQOFLATGTLRGDEKDESVEDIIEHYK 375
DB	464 LKDKVYNTQIRYQEESSQYVLLGTGLRGKEDPLVSDIIDYR 506
RESULT 3	
P85A_MOUSE	
ID	P85A_MOUSE STANDARD; PRT; 724 AA.
AC	P26450;
DT	01-AUG-1992 (rel. 23, Created)
DT	01-AUG-1992 (rel. 23, Last sequence update)
DT	15-JUN-2002 (rel. 41, Last annotation update)
DE	Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase p85-alpha subunit) (Ptdins-3-kinase p85-alpha) (PI3K).
GN	PI3KRL.
OC	Mus musculus (Mouse).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Seq	Accession	Score	Length	Matches	Mismatches	Gaps
OC	NCBI_TaxID=10090,	7.4%	149	724	17	
OX	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.	21.7%	149	724	17	
RY	[1]					
RY	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.					
RP	STRAIN-BALB/c;					
RC	MEDLINE=91191564; PubMed=1849460;					
RX	Escobedo J.A., Navankasitussas S., Kavanaugh W.M., Milfay D.,					
RA	Field V.A., Williams L.T.					
RA	"cDNA cloning of a novel 65 kd protein that has SH2 domains and					
RT	regulates binding of p13-kinase to the PDGF beta-receptor."					
RL	Cell 65:75-82(1991).					
CC	-1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE					
CC	KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING					
CC	THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.					
CC	NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE					
CC	AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.					
CC	-1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)					
CC	SUBUNTS.					
CC	-1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.					
CC	-1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.					
CC	-1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.					
CC	-----					
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration					
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -					
CC	the European Bioinformatics Institute. There are no restrictions on its					
CC	use by non-profit institutions as long as its content is in no way					
CC	modified and this statement is not removed. Usage by and for commercial					
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/					
CC	or send an email to license@isb-sib.ch).					
CC	-----					
DR	EMBL; M60651; AAA39886.1; -					
DR	HSSP; P23727; 1BF1.					
DR	MGD; MGI:97583; P1k3rl.					
DR	InterPro; IPR001720; PI3kinase_P85.					
DR	InterPro; IPR000198; RHOGAP.					
DR	InterPro; IPR000980; SH2.					
DR	InterPro; IPR001452; SH3.					
DR	Pfam; PF00017; SH2; 2.					
DR	Pfam; PF00018; SH3; 1.					
DR	Pfam; PF00620; RHOGAP; 1.					
DR	PRINTS; PR00678; PI3KINASEP85.					
DR	PRINTS; PR00401; SH2DOMAIN.					
DR	ProDom; PD000066; SH3; 1.					
DR	ProDom; PD000093; SH2; 2.					
DR	SMART; SM00324; RHOGAP; 1.					
DR	SMART; SM00252; SH2; 2.					
DR	SMART; SM00326; SH3; 1.					
DR	PROSITE; PS50001; SH2; 2.					
DR	PROSITE; PS50002; SH3; 1.					
KW	SH3 domain; SH2 domain; Repeat.					
FT	DOMAIN 3 79 SH3.					
FT	DOMAIN 333 428 SH2 1.					
FT	DOMAIN 624 718 SH2 2.					
FT	DOMAIN					
SO	SEQUENCE 724 AA; 83414 MW; DEAA66EE297CF07A CRC64;					
Query Match	7.4%; Score 149; DB 1; Length 724;					
Best Local Similarity	21.7%; Pred. No. 0.0055;					
Matches	76; Conservative 39; Mismatches 96; Indels 140; Gaps 17;					
OY	141 PLPPRPLILPKKYQPLPPEPSSR-----PLISQRTTFPEVOGMPQSISLR 188					
OY	85 PTPKRRP-----PR--PLPVAFGSSKTEADTEQQLPLPDIAEAGPADPV-APPLIKRL 136					
OY	189 DLSEV--LEAKRVPHNOKKPESTHLE-----NQNGQLPLAIISSSFTHSHSVNR 239					
OY	137 EAIKKGLGECSTLYKQTQSSNPALRQLLDCDAASVDLEMDIVHLAIAFRKYIADLPN- 195					
OY	240 DHRGGMQPCSP-----ORQCPASC-----SPHENILPKYK----- 270					
OY	196 -----PVIVVAVYVNMMSLAQELQSPEDCIIQLKLIRLPNPHQGMILQYILKHF 248					
OY	271 -----TSW----- 273					

Dd	249	KLSQSSKSNLNLARVLSIFSPVLFRPPAASDNTFHLIKALIELLISTENNERQAPALP	308
Oy	274	-RPPPP-----KSDRDVDVQHNEWYIGERSRQAVEAFMKENKOGSFLYRDCSTKSKEE	326
Rt		: :	
Dd	309	KPKPKPTVVANNSSNNMNSLDADAEWMYDDISREEVNEK-LKDITDGTFVLRAASTKMHG	367
Oy	327	PYLAVFEYENKYVNK-IRFLERNQOFAITGTLRGDEKFDSDVEIDIEHYKN	376
Rt		: :	
Dd	368	-YTLLT-----RKGGNKKLIKITHROCKTGFSDDL-----TFNSVELINHRYN	410
RESULT 4			
ID	P85A_HUMAN	STANDARD;	PRT; 724 AA.
AC	P27986:		
Dt	01-AUG-1992 (Rel. 23, Created)		
Dt	01-AUG-1992 (Rel. 23, Last sequence update)		
Dt	15-JUN-2002 (Rel. 41, Last annotation update)		
De	Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase p85-alpha subunit) (Ptdins-3-kinase p85-alpha) (PI3K).		
GN	PIK3R1 OR GRB1.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OX	Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.		
RN	NCBI_TaxID=9606;		
RN	[1]		
Rp	SEQUENCE FROM N.A.		
RX	MEDLINE=91191565; PubMed=1849461;		
RA	Skolnik E.Y., Margolis B., Mohammadi M., Lowenstein E., Fischer R.,		
RA	Drepps A., Ullrich A., Schlessinger J.,		
Rt	"Cloning of PI3 kinase-associated p85 utilizing a novel method for		
Rt	expression/cloning of target proteins for receptor tyrosine		
Rt	kinases."/;		
RL	Cell 65:83-90(1991).		
RN	[2]		
Rp	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 1-85.		
RX	MEDLINE=96196433; PubMed=8648629;		
RA	Liang J., Chen J.K., Schreiber S.L., Clardy J.;		
Rt	"Crystal structure of PI3K SH3 domain at 2.0-A resolution.";		
RL	J. Mol. Biol. 257:632-643(1996).		
RN	[3]		
Rp	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 324-434.		
RX	MEDLINE=96185448; PubMed=8599763;		
RA	Nolte R.T., Eck M.J., Schlessinger J., Shelson S.E., Harrison S.C.;		
Rt	"Crystal structure of the PI 3-Kinase p85 amino-terminal SH2 domain		
RL	and its phosphopeptide complexes.";		
RN	Nat. Struct. Biol. 3:364-373(1996).		
RN	[4]		
Rp	X-RAY CRYSTALLOGRAPHY (2.0 ANGSTROMS) OF 115-298.		
RX	MEDLINE=97121392; PubMed=8962058;		
RA	Musacchio A., Cantley L.C., Harrison S.C.;		
Rt	"Crystal structure of the breakpoint cluster region-homology domain		
RL	from Phosphoinositide 3-kinase p85 alpha subunit.";		
RN	Proc. Natl. Acad. Sci. U.S.A. 93:14373-14378(1996).		
RN	[5]		
Rp	STRUCTURE BY NMR OF 1-79.		
RX	MEDLINE=93208889; PubMed=7681364;		
RA	Koyama S., Yu H., Dalgaro D.C., Shin T.B., Zydowsky L.D.,		
RA	Schreiber S.L.;		
Rt	"Structure of the PI3K SH3 domain and analysis of the SH3 family.";		
RL	Cell 72:945-952(1993).		
RN	[6]		
Rp	STRUCTURE BY NMR OF 91-104.		
RX	MEDLINE=97121261; PubMed=8961927;		
RA	Renzoni D.A., Pugh D.J., Siligardi G., Das P., Morton C.J., Rossi C.,		
RA	Waterfield M.D., Campbell I.D., Ladbury J.E.;		
Rt	"Structural and thermodynamic characterization of the interaction of		
Rt	the SH3 domain from Fyn with the proline-rich binding site on the p85		
RL	subunit of PI3-kinase.";		
RN	Biochemistry 35:15646-15653(1996).		
RN	[7]		
Rp	STRUCTURE BY NMR OF 617-724.		
RX	MEDLINE=96312935; PubMed=86700861;		

RA Brezina A., Kara B.V., Barritt D.G., Anderson M., Smith J.C.,
RA Lake R.W., Best J.R., Cartledge S.A.:
RT "Structure of a specific peptide complex of the carboxy-terminal SH2
RT domain from the p85 alpha subunit of phosphatidylinositol 3-kinase";
RL EMBO J. 15:3579-3589(1996).
[8]
RP VARIANT ILE-326.
RX MEDLINE-97184306; PubMed-9032108;
RA Hansen T., Andersen C.B., Echwald S.M., Urhammer S.A., Clausen J.O.,
RA Vestergaard H., Owens D., Hansen L., Pedersen O.;
RT "Identification of a common amino acid polymorphism in the p85alpha
RT regulatory subunit of phosphatidylinositol 3-kinase: effects on
RT glucose disappearance constant, glucose effectiveness, and the
RT insulin sensitivity index";
RL Diabetes 46:494-501(1997).
[9]
RP VARIANT INSULIN RESISTANCE GLN-409, AND VARIANT ILE-326.
RX MEDLINE-20230645; PubMed-10768093;
RA Baynes K.C.R., Beeton C.A., Panayotou G., Stein R., Soos M.,
RA Hansen T., Simpson H., O'Rahilly S., Shepherd P.R., Whitehead J.P.;
RT "Natural variants of human p85 alpha phosphoinositide 3-kinase in
RT severe insulin resistance: a novel variant with impaired
RT insulin-stimulated lipid kinase activity";
RL Diabetologia 43:321-331(2000).
CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNITS.
CC -1- DISEASE: DEFECTS IN PIK3R1 ARE A CAUSE OF SEVERE INSULIN
CC RESISTANCE.
CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
CC -1- DATABASE: NAME=PI3K; NOTE=PI3K 1:6-12(2000);
CC WWW="http://www.ncbi.nlm.nih.gov/prow/guide/1773542685.g.htm".
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M61906; -; NOT_ANNOTATED_CDS.
DR PIR; A38748; A38748.
DR PDB; 1PKS; 31-MAY-94.
DR PDB; 1PKT; 31-MAY-94.
DR PDB; 1PHT; 07-DEC-95.
DR PDB; 1PBW; 12-MAR-97.
DR PDB; 1PIC; 17-SEP-97.
DR PDB; 1AON; 25-FEB-98.
DR PDB; 1AZG; 25-FEB-98.
DR Genew; HGNC:8979; PIK3R1.
DR MIM; 171833.
DR InterPro; IPR001720; PI3Klnase_P85.
DR InterPro; IPR000198; RHO GAP.
DR InterPro; IPR000980; SH2.
DR InterPro; IPR001452; SH3.
DR Pfam; PF00017; SH2; 2.
DR Pfam; PF00018; SH3; 1.
DR Pfam; PF00620; RHO GAP; 1.
DR PRINTS; PR00678; PI3KINASEP85.
DR PRINTS; PR00401; SH2DOMAIN.
DR PRODOM; PD000066; SH3; 1.
DR PRODOM; PD000093; SH2; 2.
DR SMART; SM00334; RHO GAP; 1.
DR SMART; SM00252; SH2; 2.
DR SMART; SM00326; SH3; 1.
DR PROSITE; PS50001; SH2; 2.

RW	PROSITE PS50002; SH3; 1.
KR	SH3 domain; SH2 domain; Repeat; 3D-structure; Polymorphism;
KM	Disease mutation.
FT	DOMAIN 79 SH3.
FT	DOMAIN 333 SH2 1.
FT	DOMAIN 624 SH2 2.
FT	VARIANT 326 M -> I.
FT	/FtId=VAR_010023.
FT	Variant 409 409 R -> O (IN SEVERE INSULIN RESISTANCE;
FT	REDUCTION OF INSULIN-STIMULATED
FT	ACTIVITY).
FT	/FtId=VAR_010024.
SQ	SEQUENCE 724 AA; 83597 MM; 5DP9493C0431C76DD CRC64;
	Query Match 7.3%; Score 146.5; DB 1; Length 724;
	Best Local Similarity 24.7%. Pred. No. 0.0082;
	Matches 58; Conservative 35; Mismatches 73; Indels 69; Gaps 12;
QY	176 PEVQGMPQSILSRDLSEVLAEKYPVHNO-----RKREPTHLENTQDEI 220
	: : : : : : : : :
Db	211 PEGVSSEYIQ-LKKLIIRSPSTIHQWLFLQYLKHFPLKSQSSKNLLNARLSELF 268
QY	221 -PLAISSSSFTTSM-----SYQNDRHGGMPCSPQRCPAPASCSPHENILPY 268
	: : : : : : : : : : : : : : :
Db	269 SPMLERFPAASSDNTENLIKVELISIEWEMERQ-----PALALP----- 309
QY	269 KYTSNRPPPK-----SDRDQVNHMEYIGEYSROAVEAFEMKNGSFLYDCSTK 322
	::::: : : : : : : : : : : : : : :
Db	310 -----KPKPKPTTVANGANNNNNSLONAEHYMGDISREVENEK-LRDYADGTFIVADAETK 363
QY	323 SKPEPVLAVFENKYVNVK-TRPLENQOQALGTGLRGDEKEVDIEIHKYN 376
	: : : : : : : : : : : : : : :
Db	364 MHGD-YTITLT--RRKGNNKLIIKIFHRDGKGYFSDDL---TFSSVELLNHYRN 410
	:
RESULT 5	
P85A_BOVIN	STANDARD; PRT; 724 AA.
ID	P85A_BOVIN
AC	P23727;
DT	01-NOV-1991 (Rel. 20, Created)
DT	01-NOV-1991 (Rel. 20, Last sequence update)
DT	15-JUN-2002 (Rel. 41, Last annotation update)
DE	Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase p85-alpha subunit) (Ptdins-3-kinase p85-alpha) (PI3K).
DE	p85.alpha subunit) (Ptdins-3-kinase p85-alpha) (PI3K).
GN	PI3Krl.
OS	Bos taurus (Bovine).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
CC	Bovidae; Bos.
OX	NCL_TaxID=9913;
XM	[1]
RN	SEQUENCE FROM N.A.
RP	MEDLINE=91191567; PubMed=1707345;
RA	Otsu M., Hiles I.D., Gout I., Fry M.J., Ruiz-Larrea F., Panayotou G.,
RA	Thompson A., Dhand R., Hsuan J., Totty N., Smith A.D., Morgan S.J.,
RA	Courtenidge S.A., Parker P.J., Waterfield M.D.:
RT	"Characterization of two 85 kd proteins that associate with receptor
RT	tyrosine kinases, middle-T/pp60c-src complexes, and PI3-kinase.";
RL	Cell 65:91-104(1991).
RM	[2]
RN	CIRCULAR DICHOISM AND FLUORESCENCE SPECTROSCOPY.
RP	MEDLINE=93049176; PubMed=1330535;
RA	Panayotou G., Bax B., Gout I., Federwisch M., Wroblewski B., Dhand R.,
RA	Fry M.J., Bindell T.L., Moliner A., Waterfield M.D.:
RT	"Interaction of the p85 subunit of PI 3-kinase and its N-terminal SH2
RT	domain with a PDF receptor phosphorylation site: structural features
RT	and analysis of conformational changes.";
RT	EMBO J. 11:4261-4272(1992).
RN	[3]
RP	STRUCTURE BY NMR OF 1-84.
RA	MEDLINE=93272320; PubMed=7684655;
RA	Booker G.W., Gout I., Downing A.K., Driscoll P.C., Boyd J.,
RA	Waterfield M.D., Campbell I.D.;

RT "Solution structure and ligand-binding site of the SH3 domain of the
 RT p85 alpha subunit of phosphatidylinositol 3-kinase.";
 RL Cell 73:813-822(1993).
 RN [4]
 RE STRUCTURE BY NMR OF 314-431.
 RX MEDLINE=92357146; PubMed=1323062;
 RA Booker G.W., Breeze A.L., Downing A.K., Panayotou G., Gout I.,
 RA Waterfield M.D., Campbell I.D.;
 RT "Structure of an SH2 domain of the p85 alpha subunit of
 RT phosphatidylinositol-3-OH kinase.";
 RL Nature 358:684-687(1992).
 RN [5]
 RE STRUCTURE BY NMR OF 321-434.
 RX MEDLINE=97110350; PubMed=8952511;
 RA Guenther U.L., Liu Y., Sanford D., Bachovich W.W., Schaffhausen B.;
 RT "NMR analysis of interactions of a phosphatidylinositol 3'-kinase SH2
 RT domain with phosphotyrosine peptides reveals interdependence of major
 RT binding sites.";
 RL Biochemistry 35:15570-15581(1996).
 RN [6]
 RE STRUCTURE BY NMR OF 614-724.
 RX MEDLINE=98173872; PubMed=9512716;
 RA Siegel G., Davis B., Kristensen S.M., Sankar A., Linacre J.,
 RA Stein R.C., Panayotou G., Waterfield M.D., Driscoll P.C.;
 RT "Solution structure of the C-terminal SH2 domain of the p85 alpha
 RT regulatory subunit of phosphoinositide 3-kinase.";
 RL J. Mol. Biol. 276:461-478(1998).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES. THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC NECESSARY FOR THE INSULIN-STIMULATED INCREASE IN GLUCOSE UPTAKE
 CC AND GLYCOGEN SYNTHESIS IN INSULIN-SENSITIVE TISSUES.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNITS.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M61745; AAA79511.1; -
 DR PIR: A38749; A38749.
 DR PDB: 2PNA; 31-JAN-94.
 DR PDB: 2PNI; 31-JAN-94.
 DR PDB: 1PNJ; 31-OCT-93.
 DR PDB: 1BFI; 25-FEB-98.
 DR PDB: 1BFI; 25-FEB-98.
 DR InterPro: IPR001720; PI3K_kinase_p85.
 DR InterPro: IPR00198; RhoGAP.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhoGAP; 1.
 DR PRINTS: PR00478; PI3K_kinase_p85.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RhoGAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 KW SH3 domain; SH2 domain; Repeat; 3D-structure.
 FT DOMAIN 3 79
 FT SH3.
 FT RC TISSUE=Liver;
 FT RX MEDLINE=97218222; PubMed=9065454;
 FT DOMAIN 129 273
 FT GAP DOMAIN.

FT DOMAIN 333 428 SH2 1.
 FT DOMAIN 624 718 SH2 2.
 FT HELIX 340 347
 FT TURN 348 348
 FT STRAND 354 359
 FT STRAND 354 359
 FT STRAND 368 373
 FT STRAND 378 382
 FT STRAND 384 384
 FT STRAND 391 391
 FT STRAND 401 409
 FT HELIX 401 413
 FT STRAND 413 413
 FT TURN 418 419
 FT STRAND 422 422
 FT STRAND 427 427
 SQ SEQUENCE 724 AA; 83497 MM; EDDF6E754BBE7321 CMC64;
 Query Match 7.2%; Score 146; DB 1; Length 724;
 Best Local Similarity 22.4%; Pred. No. 0.0088;
 Matches 77; Conservative 49; Mismatches 103; Indels 114; Gaps 16;
 Oy 136 RKNKTPLEPPRLITLPKKYQPLPEPPSSR-----PLSQQRHTPEV----- 178
 Db 80 KKSPPTRKPPRPRLPVAPGSKTEADSEQASTLPDLADQFAPPDVAPPLIKLYEAI 139
 Oy 179 --QGMPQSIQR-----DLSEVLEAEKVP-----HNQRKESTHLENTQEI 220
 Db 140 EKKGLECSTLYRTQSSSNPAELRLDDCDTASLDLDEMDEVHLADAFKRYLLDLPN-PVI 198
 Oy 221 PLAISSSFT----- 236
 Db 199 PVAASSEISLAPEVQSSSEYIQLKKLRSPSPHQYWLTLQYLLKHFKLQSSKNL 258
 Oy 227 ONRDHGMOPQPCORCPASCSPH-ENILP-----YKTSR-----PPPKR--- 280
 Db 259 LNA--RVISELFPLEFPAPASSENTHLIKITILISTENENQPPALDPPKPKPTT 316
 Oy 281 -----SDRKDVQHNEMYIGEYSRQAVEAFPKENKDSFLVRCSTFSKEEYVLAIFY 334
 Db 317 VANNGMNNMNSLDQDEAWGDISREVENK-LRDTADGTPLVROASTMHHQD-YLLTL-- 372
 Oy 335 EKKYVNVK-IRPLENQOPALCTGRCDEKPDVSDEIIEHKY 376
 Db 373 -RKGNMNLKIFHRDGRYGFSDPL---TFNSVVELINHYRN 410
 RESULT 6
 P85A_RAT STANDARD; PRT: 724 AA.
 AC Q63787; Q63790; P70544; O55085;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory alpha subunit (PI3-kinase
 DE p85-alpha subunit) (PtdIns-3-kinase p85-alpha) (PI3K).
 GS PIR31.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS P85-ALPHA AND P55-ALPHA).
 RC STRAIN=Wistar; TISSUE=Brain;
 RX MEDLINE=96214979; PubMed=862182;
 RA Inukai K., Anai M., van Breda E., Hosaka T., Katagiri H., Funaki M.,
 RA Fukushima Y., Ogihara T., Yazaki Y., Kikuchi M., Oka Y., Asano T.;
 RT "A novel 55-kDa regulatory subunit for phosphatidylinositol 3-kinase
 RT of the p85alpha gene.";
 RL J. Biol. Chem. 271:5317-5320(1996).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM P50-ALPHA).
 RC TISSUE=Liver;
 RX MEDLINE=97218222; PubMed=9065454;

RT tyrosine kinase (Btk) loci.";
 RL J. Immunol. 153:5607-5617(1994).
 [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C129;
 RX MEDLINE=95352959; PubMed=7626884;
 RA Oeljen J.C., Liu X., Lu J., Allen R.C., Muzny D.M., Belmont J.W.,
 RA Gibbs R.A.;
 RT "Sixty-nine kilobases of contiguous human genomic sequence containing
 RT the alpha-galactosidase A and Bruton's tyrosine kinase loci.";
 RL Mamm. Genome 6:334-338(1995).
 [5]
 RP VARIANT XID CYS-28.
 RX MEDLINE=93324903; PubMed=8332901;
 RA Rawlings D.J., Saffran D.C., Tsukada S., Largaespada D.A.,
 RA Grimaldi J.C., Cohen L., Mohr R.N., Bazan J.F., Howard M.,
 RA Copeland N.G., Jenkins N.A., White O.N.;
 RT "Mutation of unique region of Bruton's tyrosine kinase in
 RT immunodeficient Xid mice.";
 RL Science 261:358-361(1993).
 CC -1- FUNCTION: PLAYS A CRUCIAL ROLE IN B-CELL ONTOGENY.
 CC -1- CATALYTIC ACTIVITY: ATP + a protein tyrosine - ADP + protein
 CC tyrosine phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic and membrane-associated.
 CC -1- PTM: Autophosphorylated on Tyr-223 and Tyr-551. The tyrosine
 CC phosphorylation of Tyr-223 may create a docking site for a SH2
 CC containing protein.
 CC -1- DISEASE: DEFECTS IN BTK ARE THE CAUSE OF MURINE X-LINKED
 CC IMMUNODEFICIENCY (XID).
 CC -1- SIMILARITY: BELONGS TO THE TYR FAMILY OF PROTEIN KINASES. BTK
 CC SUBFAMILY.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC
 DR EMBL: L08967; AAA37316.1; -;
 DR EMBL: L10627; -; NOT_ANNOTATED_CDS.
 DR EMBL: L29788; AAA66943.1; -;
 DR EMBL: U58105; AAB47246.1; -;
 DR HSSP: Q06187; IBMN.
 DR MGD: MGI:88216; Btk.
 DR InterPro: IPR001562; BTK.
 DR InterPro: IPR000719; Euk_pkinase.
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR InterPro: IPR001245; Tyr_pkinase.
 DR Pfam: PF00017; SH2; 1.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00069; pkinase; 1.
 DR Pfam: PF00169; PH; 1.
 DR Pfam: PF00779; BTK; 1.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PRINTS: PR00402; TECTKINASE.
 DR PRINTS: PR00109; TYRKINASE.
 DR ProDom: PD000001; Euk_pkinase; 1.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 1.
 DR SMART: SM00107; BTK; 1.
 DR SMART: SM00233; PH; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00326; SH3; 1.
 DR SMART: SM00219; Tyrc; 1.
 DR PROSITE: PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE: PS00109; PROTEIN_KINASE_TYR; 1.

DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE: PS50001; SH2; 1.
 DR PROSITE: PS50002; SH3; 1.
 DR PROSITE: PS50003; PH_DOMAIN; 1.
 DR Transferrase; Tyrosine-protein kinase; Phosphorylation;
 KW ATP-binding; SH3 domain; SH2 domain; Disease mutation;
 FT DOMAIN 3 133
 FT DOMAIN 214 274
 FT DOMAIN 281 377
 FT DOMAIN 402 655
 FT NP_BIND 408 416
 FT BINDING 430 430
 FT ACT_SITE 521 521
 FT MOD_RES 223 223
 FT MOD_RES 551 551
 FT VARIANT 28 28
 FT MUTAGEN 41 41
 FT MUTAGEN 223 223
 FT MUTAGEN 430 430
 FT CONFLICT 67 67
 FT CONFLICT 123 123
 FT CONFLICT 197 197
 FT CONFLICT 450 450
 SQ SEQUENCE 659 AA; 76437 MW; E502B798BC36E223 CRC64;
 Query Match 7.1%; Score 143; DB 1; Length 659;
 Best Local Similarity 23.4%; Pred No. 0.013; Indels 98; Gaps 11;
 Matches 60; Conservative 27; Mismatches 71;
 QY 125 RSONKGASVRRKKIPLPPRPRLTLPRKYOPLPEPPESSRPPLSQRRHTEPVQGMPSQ 184
 DB 171 RNSGLKPGSSHRKTRKPLPTPEEDQILKK--PLPEPTA-----Ap 210
 QY 185 ISRLDSEVLEA-EKYPHNO-----RKPESTHLENQOIEPLAISSSFTTNSVON 228
 DB 211 ISTTELRKVALDYVPMNANDQLKRGGEYFLTESN--LFW-----WRA 254
 QY 239 RDRHGMQCSPPQRCPPASCSPHENILPKYTSWMPRPKRSDRKDVQHNEMYIGEYSR 298
 DB 255 RDKNG-----QEGYIPSNYIT-----EAEDSIENYEMYSKHMTR 288
 QY 299 QAEEAFMKENKDGSLVNDCKSTKSEEPYVLAIFYENKYNKIRFLERNOQFALGTGL 358
 DB 289 SQAEQLKQEGKRGCFIVDSSKAGK-----YTVSV-----FANSTG- 325
 QY 359 RGDEKFDVSDITIEHY 374
 DB 326 -----EPQGVIRHY 334
 RESULT 8
 3BP2_MOUSE
 ID 3BP2_MOUSE STANDARD; PRT; 559 AA.
 AC 006649;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE SH3 domain-binding protein 2 (3BP-2).
 GN SH3BP2 OR 3BP2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93174278; PubMed=8438166;
 RA Ren R., Mayer B.J., Cicchetti P., Baltimore D.;
 RT "Identification of a ten-amino acid proline-rich SH3 binding site.";
 RL Science 259:1157-1161(1993).
 CC -1- FUNCTION: BINDS DIFFERENTIALLY TO THE SH3 DOMAINS OF CERTAIN
 CC PROTEINS OF SIGNAL TRANSDUCTION PATHWAYS. BINDS TO
 CC PHOSPHATIDYLINOSITOLS; LINKING THE HEMOPOIETIC TYROSINE KINASE

FES TO THE CYTOPLASMIC MEMBRANE IN A PHOSPHORYLATION DEPENDENT MECHANISM (BY SIMILARITY).

-1- SIMILARITY: CONTAINS 1 PH DOMAIN.

-1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sid.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; L14543; AAA37121.1; .

DR MGD; MG1;1346349; Sh3bp2.

DR InterPro: IPR001849; PH.

DR InterPro: IPR000980; SH2.

DR Pfam: PF00017; SH2; 1.

DR Pfam: PF00169; PH; 1.

DR ProDom: PD000093; SH2; 1.

DR SMART; SM00233; PH; 1.

DR SMART; SM00252; SH2; 1.

DR PROSITE; PS50001; SH2; 1.

DR PROSITE; PS50003; PH;DOMAIN; 1.

DR SH2 domain; SH3-binding.

KW DOMAIN 26 130 PH.

FT SITE 201 210 SH3-BINDING.

FT DOMAIN 205 212 POLY-PRO.

FT DOMAIN 236 240 POLY-PRO.

FT FT 455 553 SH2.

SEQUENCE 559 AA; 62208 MW; EDPEFL1B259646E CRC64;

Query Match 6.9%; Score 140; DB 1; Length 559;

Best Local Similarity 19.7%; Pred. No. 0.017;

Matches 93; Conservative 54; Mismatches 156; Indels 168; Gaps 18;

QY 56 LRHEETWQSKILIPARIKESEYADTHYFKVAND-----TPPLD 95

DB 82 MRAAETTSNNVFPFKIHISKRRHWFESASEDERKSMATVRRRIGFHEKKELPLD 141

QY 96 TRTSISIGPTMTQRFLEVDKPISRDVSONIKGDASVKKKIP--LRPPR----- 147

DB 142 TSDSSS-----DTDSYGAVERPIDISLSSYPMDNEDYHEDEDSYLEPDSFGPKLE 195

QY 148 -LITLKKIOPLRPEPSSRPPLS--QRHTF-----PEVQMPQSISL-RDL 190

DB 196 DALTYPPAYPP-PPVPPYRKPAFSDLPRAHSFTSKSPPLPPPPKRGLPDGSAPEDA 254

QY 191 SEVLEAEKVNHNORKEPSTHLLNOMTQELPLAIS----- 225

DB 255 KDALGLRRVERGLRVPATPRRMSDPKSNVPTVNLKRNKCFRDSVNPGLPWTGHTS 314

QY 226 -SSSEFTSNHVSQVNRD-----HRGGMQ-----CSFORQ--- 254

DB 315 SVSSSTTMVAATSRNCDKLKSFLLSSRGPTSPRPYPAKPKFKLIAEPSTREAKKFA 374

QY 255 --PPASCSPEHNTLPYKYSWRPPFKRSRDKVQH----- 288

DB 375 PVPPVAPRPVQKMPAEAVBRAPVLRPRENTPLPHLQSRPDGQSPRGSEFEKARQPSQ 434

QY 289 -----NEMTYGEYSKRAVEAEAFMK-----ENKQGSFLVRCQSKRSKE 335

DB 435 ADTGEDESDDEYKVPKPSNVFNTTESCEVERLFRATDPGRGPODGLYCIIRNSTSKGK 494

QY 326 EPYVLAVEYE--NKVYNVNRIRFLERNQOFALGTGLGDEKDEKDSVEDIIEHY 374

DB 495 ----VLVVMDESSKKVKNYRL--FEKDSKFT-----LGEVLFASVGSKEVHT 536

DT 01-JUN-1994 (Rel. 29, Created)

DT 15-JUN-1994 (Rel. 29, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tyrosine-protein kinase btk (EC 2.7.1.112) (Bruton's tyrosine kinase)

DE (Agammaglobulinemia tyrosine kinase) (ATK) (B cell progenitor kinase) (BPK).

GN BTK OR ATK OR AGMX1 OR BPK.

OS Homo sapiens (Human).

OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

NCBI_TaxID=9606;

OX [1]

RN SEQUENCE FROM N.A.

RP MEDLINE=93140868; PubMed=8380905;

RX Vetric D., Vorechovsky I., Sideras P., Holland J., Davies A.,

RA Flinter F., Hammarstrom L., Kinnon C., Levinsky R.J., Bobrow M.,

RA Smith C.I.E., Bentley D.R.;

RT "The gene involved in X-linked agammaglobulinemia is a member of the

RL src family of protein-tyrosine kinases.";

RL Nature 361:226-233(1993).

[2]

RN ERRATUM.

RP Vetric D., Vorechovsky I., Sideras P., Holland J., Davies A.,

RA Flinter F., Hammarstrom L., Kinnon C., Levinsky R.J., Bobrow M.,

RA Smith C.I.E., Bentley D.R.;

RL Nature 364:362-362(1993).

[3]

RN SEQUENCE FROM N.A.

RP TISSUE=Blood;

RX MEDLINE=943377492; PubMed=8090769;

RA Ohta Y., Haire R.N., Litman R.T., Fu S.M., Nelson R.P., Kratz J.,

RA Kornfeld S.J., La Morena M., Good R.A., Litman G.W.;

RT "Genomic organization and structure of Bruton agammaglobulinemia

RT tyrosine kinase: localization of mutations associated with varied

RT clinical presentations and course in X chromosome-linked

RT agammaglobulinemia.";

RL Proc. Natl. Acad. Sci. U.S.A. 91:9062-9066(1994).

[4]

RN SEQUENCE FROM N.A.

RP MEDLINE=95012452; PubMed=7927535;

RX Rohrer J., Parolini O., Belmont J.W., Conley M.E.;

RT "The genomic structure of human BTK, the defective gene in X-linked

RT agammaglobulinemia.";

RL Immunogenetics 40:319-324(1994).

[5]

RN SEQUENCE FROM N.A.

RP MEDLINE=95352959; PubMed=7626884;

RX Oeltjen J.C., Liu X., Lu J., Allen R.C., Muzny D.M., Belmont J.W.,

RA Gibbs R.A.;

RT "Sixty-nine kilobases of contiguous human genomic sequence containing

RT the alpha-galactosidase A and Bruton's tyrosine kinase loci.";

RL Mamm. Genome 6:334-338(1995).

[6]

RN SEQUENCE FROM N.A.

RP Oeltjen J.C., Malley T.M., Muzny D.M., Miller W., Gibbs R.A.,

RA Belmont J.W.;

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.

[7]

RN SEQUENCE OF 1-442 FROM N.A.

RP MEDLINE=93145329; PubMed=8425221;

RX Tsukada S., Saffian D.C., Rawlings D.J., Parolini O., Allen R.C.,

RA Kilsak I., Sparks R.S., Kubagawa H., Mohandas T., Quan S.,

RA Belmont J.W., Cooper M.D., Conley M.E., Witte O.N.;

RT "Deficient expression of a B cell cytoplasmic tyrosine kinase in

RT human X-linked agammaglobulinemia.";

RL Cell 72:279-290(1993).

[8]

RN PHOSPHORYLATION OF GTF2I, AND MUTAGENESIS.

RP MEDLINE=97165069; PubMed=9012831;

RX Yang W., Desiderio S.;

RT "BAP-135, a target for Bruton's tyrosine kinase in response to B cell

RT receptor engagement.";

RL Proc. Natl. Acad. Sci. U.S.A. 94:604-609(1997).

[9] X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 1-170.
 RP MEDLINE-97361821; PubMed-9218782;
 RA Hyonenen M., Saraste M.;
 RT "Structure of the PH domain and Btk motif from Bruton's tyrosine
 kinase: molecular explanations for X-linked agammaglobulinemia.";
 RL EMBO J. 16:3396-3404(1997).
 RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-170.
 RP MEDLINE-99216540; PubMed-10196129;
 RA Baraldi E., Carrugo K.D., Hyonenen M., Surdo P.L., Riley A.M.,
 RA Potter B.V.L., O'Brien R., Laddbury J.E., Saraste M.;
 RT "Structure of the PH domain from Bruton's tyrosine kinase in complex
 with inositol 1,3,4,5-tetrakisphosphate.";
 RL Structure 7:449-460(1999).
 RP STRUCTURE BY NMR OF 209-275.
 RP MEDLINE-98153176; PubMed-9485443;
 RA Hansson H., Mattsson P.T., Allard P., Haapaniemi P., Vihinen M.,
 RA Smith C.I.E., Haerd T.;
 RT "Solution structure of the SH3 domain from Bruton's tyrosine kinase.";
 RL Biochemistry 37:2912-2924(1998).
 RP REVIEW ON XLA VARIANTS.
 RP MEDLINE-96174621; PubMed-8594569;
 RA Vihinen M., Iwata T., Kinnon C., Kwan S.-P., Ochs H.D.,
 RA Vorechovsky I., Smith C.I.E.;
 RT "Btkbase, mutation database for X-linked agammaglobulinemia (XLA).";
 RL Nucleic Acids Res. 24:160-165(1996).
 RP REVIEW ON XLA VARIANTS.
 RP MEDLINE-97169387; PubMed-9016530;
 RA Vihinen M., Belohradsky B.H., Haire R.N., Holinski-Feder E.,
 RA Kwan S.-P., Lappalainen I., Lehtvaslahti H., Lester T., Meindl A.,
 RA Ochs H.D., Ollila J., Vorechovsky I., Weiss M., Smith C.I.E.,
 RT "Btkbase, mutation database for X-linked agammaglobulinemia (XLA).";
 RL Nucleic Acids Res. 25:166-171(1997).
 RP VARIANTS XLA TRP-288; GLY-307; ASP-607 AND S-V-F-S-T-R-103 INS.
 RP MEDLINE-94214481; PubMed-8162056;
 RA Bradley L.A.D., Sweetman A.K., Lovering R.C., Jones A.M., Morgan G.,
 RA Levinsky R.J., Kinnon C.;
 RT "Mutation detection in the X-linked agammaglobulinemia gene, BTK,
 RT using single strand conformation polymorphism analysis.";
 RL Hum. Mol. Genet. 3:79-83(1994).
 RP VARIANTS XLA HIS-28 AND TRP-288.
 RP MEDLINE-94214435; PubMed-8162018;
 RA de Weers M., Mensink R.G.J., Kraakman M.E.M., Schuurman R.K.B.,
 RA Hendriks R.W.;
 RT "Mutation analysis of the Bruton's tyrosine kinase gene in X-linked
 RT agammaglobulinemia: identification of a mutation which affects the
 RT same codon as is altered in immunodeficient xid mice.";
 RL Hum. Mol. Genet. 3:161-166(1994).
 RP SEQUENCE FROM N.A., AND VARIANTS XLA S-334; R-506; Q-520; W-562 AND
 RP K-630.
 RP MEDLINE-95152493; PubMed-7880320;
 RA Hagemann T.L., Chen Y., Rosen F.S., Kwan S.-P.;
 RT "Genomic organization of the Btk gene and exon scanning for mutations
 RT in patients with X-linked agammaglobulinemia.";
 RL Hum. Mol. Genet. 3:1743-1749(1994).
 RP VARIANTS XLA D-113; C-361; Q-520; P-542; W-562; K-630 AND P-652.
 RP MEDLINE-95152494; PubMed-7849697;
 RA Conley M.E., Fitch-Hilgenberg M.E., Cleveland J.L., Parolini O.,
 RA Rohrer J.;
 RT "Screening of genomic DNA to identify mutations in the gene for
 RT Bruton's tyrosine kinase.";
 RL Hum. Mol. Genet. 3:1751-1756(1994).
 RP VARIANTS XLA H-28; P-33; P-408; G-589; D-613 AND 260-Q--E-280 DEL.
 RP MEDLINE-95152522; PubMed-7849721;

RA Zhu Q., Zhang M., Winkelstein J., Chen S.-H., Ochs H.D.;
 RT "Unique mutations of Bruton's tyrosine kinase in fourteen unrelated
 RT X-linked agammaglobulinemia families.";
 RL Hum. Mol. Genet. 3:1899-1900(1994).
 RP VARIANTS XLA E-430; Q-520; Q-525; P-562; V-582; G-589; E-594 AND
 RP D-613.
 RP MEDLINE-95108046; PubMed-7809124;
 RA Vihinen M., Vetrle D., Maniar H.S., Ochs H.D., Zhu Q., Vorechovsky I.,
 RA Webster A.D.B., Notarangelo L.D., Nilsson L., Sowatski J.M.,
 RA Smith C.I.E.;
 RT "Structural basis for chromosome X-linked agammaglobulinemia: a
 RT tyrosine kinase disease.";
 RL Proc. Natl. Acad. Sci. U.S.A. 91:12803-12807(1994).
 RP VARIANT XLA PHE-64, AND CHARACTERIZATION OF OTHER XLA VARIANTS.
 RP MEDLINE-95151728; PubMed-7849006;
 RA Vihinen M., Zvelebil J.J.M., Zhu Q., Brocimans R.A., Ochs H.D.,
 RA Zegers B.J.M., Nilsson L., Waterfield M.D., Smith C.I.E.;
 RT "Structural basis for pleckstrin homology domain mutations in
 RT X-linked agammaglobulinemia.";
 RL Biochemistry 34:1475-1481(1995).
 RP VARIANTS XLA S-25; W-288; M-370; V-509; P-525; K-526; W-562
 RP AND R-594.
 RP MEDLINE-95227177; PubMed-7711734;
 RA Vorechovsky I., Vihinen M., de Saint Basile G., Honsova S.,
 RA Hammarstrom L., Mueller S., Nilsson L., Fischer A., Smith C.I.E.;
 RT "DNA-based mutation analysis of Bruton's tyrosine kinase gene in
 RT patients with X-linked agammaglobulinemia.";
 RL Hum. Mol. Genet. 4:51-58(1995).
 RP VARIANTS XLA LYS-567; LEU-587 AND HIS-641.
 RP MEDLINE-95359977; PubMed-7633420;
 RA Jin H., Webster A.D.B., Vihinen M., Sideras P., Vorechovsky I.,
 RA Hammarstrom L., Bernatowska-Matuszkiewicz E., Smith C.I.E.,
 RA Bobrow M., Vetrle D.;
 RT "Identification of Btk mutations in 20 unrelated patients with
 RT X-linked agammaglobulinemia (XLA).";
 RL Hum. Mol. Genet. 4:693-700(1995).
 RP VARIANTS XLA PRO-33; GLN-520; CYS-641 AND GLY-302 DEL.
 RP MEDLINE-95359987; PubMed-7633429;
 RA Gaspar H.B., Bradley L.A.D., Katz F., Lovering R.C., Rolfman C.M.,
 RA Morgan G., Levinsky R.J., Kinnon C.;
 RT "Mutation analysis in Bruton's tyrosine kinase, the X-linked
 RT agammaglobulinemia gene, including identification of an insertional
 RT hotspot.";
 RL Hum. Mol. Genet. 4:755-757(1995).
 RP VARIANTS XLA ASN-429 AND ARG-477.
 RP MEDLINE-96177680; PubMed-8634718;
 RX
 Query Match 6.9%; Score 140; DB 1; Length 659;
 Best local Similarity 22.7%; Pred. No. 0.02;
 Matches 58; Conservative 29; Mismatches 71; Indels 98; Gaps 11;
 QY 125 RQNIINGDASVRRKNTPLPPRPLILPKKYQLPPPESSRRPLSQRHTFFVQGMPSQ 184
 DB 171 RNSLSLPGSSHRKTRPLPTPEEDQILKK--PLPEPPA-----Ap 210
 QY 185 ISLRDSEVLVA-EKYPHND-----RKPESTHLENONQEIPLAISSSFTSNHNVON 238
 DB 211 VTSSELKRVVALLDYPMNANDQLKRGDEYFLLESN---LPW-----WRA 254
 QY 239 RDHRGGMQPCPCPCPPASCSPHENILPYKYTSWRRPPKRSRDKRDVOHNEWYIGEYSR 298
 DB 255 RDKNG-----QEGYIPSNVY-----EAEDSIEMWYSKHMTR 288
 QY 299 QAVEEAPFKENKGSFLVDCSTKSKPEPYVLAVFENKYNKIKFLENOOPALGTGL 358
 DB 289 SOAEOLLKQEGKGGFIVRDSKAGK-----YTVSV-----FAKSTG- 325

Qy 359 RGDGPDSDVEDIEHY 374
 Db 326 -----DPOGVIRHY 334

RESULT 10

CSW_DROME STANDARD; PRT: 845 AA.
 ID CSW_DROME
 AC P29349: Q24032: Q9W524: Q24033: Q9V3H1:
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Protein-tyrosine phosphatase corkscrew (EC 3.1.3.48).
 GN CSM OR EG:BACN25G24.2 OR CG3954.
 OS Drosophila melanogaster (Fruit fly).
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RP SEQUENCE FROM N.A. (ISOFORM Y1229), FUNCTION, TISSUE SPECIFICITY, AND
 RP DEVELOPMENTAL STAGE.
 RC TISSUE-Embryo;
 RX MEDLINE=92346711; PubMed=1638629;
 RA Perkins L.A., Larsen I., Perrimon N.;
 RT "Corkscrew encodes a putative protein tyrosine phosphatase that
 RT functions to transduce the terminal signal from the receptor tyrosine
 RT kinase torso."
 RL Cell 70:225-236(1992).
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS Y1229 AND 4A).
 RC STRAIN-DP CN BW;
 RA Melnick M.B., Melnick C.B., Larsen I., Perrimon N., Perkins L.A.;
 RT "The role of the Drosophila corkscrew protein as a transducer
 RT downstream of receptor tyrosine kinases is functionally conserved."
 RT Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORMS Y1229; 2 AND 4A).
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,
 RA Wen R.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista A.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Fosler K., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kemison J.A., Kiechallum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasok P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Murray D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORMS 4A AND Y1229).
 RC STRAIN-Oregon-R;
 RX MEDLINE=20196011; PubMed=10731137;
 RA Benos P.V., Gatt M.K., Ashburner M., Murphy L., Harris D.,
 RA Bartell B.G., Ferraz C., Vidal S., Brun C., Demallies J., Cadieu E.,
 RA Dreano S., Gloux S., Lelaure V., Mottier S., Galibert F., Borokova D.,
 RA Minana B., Kafatos F.C., Louis C., Siden-Kiamos I., Bolshakov S.,
 RA Papagiannakis G., Spanos L., Cox S., Medueno E., de Pablo B.,
 RA Papadogiannakis J., Peter A., Schoettler P., Werner M., Mourikoti F.,
 RA Benet N., Dove G., Schaefer U., Jockle H., Bucheton A.,
 RA Callister D.M., Campbell L.A., Darlamiou A., Henderson N.S.,
 RA McMillan P.J., Sallies C., Tait E.A., Valenti P., Saunders R.D.C.,
 RA Glover D.M.;
 RT "From sequence to chromosome: the tip of the X chromosome of D.
 RT melanogaster."
 RL Science 287:2220-2222(2000).
 RN [5]
 RP FUNCTION.
 RX MEDLINE=97105827; PubMed=8948575;
 RA Perkins L.A., Johnson M.R., Melnick M.B., Perrimon N.;
 RT "The nonreceptor protein tyrosine phosphatase corkscrew functions in
 RT multiple receptor tyrosine kinase pathways in Drosophila."
 RL Dev. Biol. 180:63-81(1996).
 CC -1- FUNCTION: REQUIRED IN ALL RECEPTOR TYROSINE KINASE SIGNALING
 CC PATHWAYS. FUNCTIONS DOWNSTREAM OF THE RECEPTOR TYROSINE KINASE
 CC TORSO, ACTING IN CONCERT WITH D-RAF VIA TAILLESS. ALSO FUNCTIONS
 CC DOWNSTREAM OF EGFR (EPIDERMAL GROWTH FACTOR RECEPTOR) AND BTL
 CC (PIRHOBLAST GROWTH FACTOR RECEPTOR). THE SH2 DOMAIN SUGGESTS THAT
 CC CSM EFFECTS ITS ROLE BY MEDIATING HETEROERIC PROTEIN
 CC INTERACTIONS. MATERINALLY REQUIRED FOR NORMAL DETERMINATION OF CELL
 CC FATES AT THE TERMINI OF THE EMBRYO. REQUIRED FOR CELL FATE
 CC SPECIFICATION OF THE VENTRAL ECTODERM, IN THE DEVELOPING EMBRYONIC
 CC CNS AND FOR EMBRYONIC TRACHEAL CELL MIGRATION. FUNCTIONS DURING
 CC IMAGINAL DEVELOPMENT FOR PROPER FORMATION OF ADULT STRUCTURES SUCH
 CC AS EYES, ARISTAE, US WING VEIN AND THE TARSAL CLAW.
 CC -1- CATALYTIC ACTIVITY: Protein tyrosine phosphatase + H(2)O = protein
 CC tyrosine + phosphate.
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS: 2 (SHOWN HERE), 4A AND Y1229;
 CC MAY BE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EXPRESSED UNIFORMLY THROUGHOUT ALL TISSUES
 CC DURING EMBRYOGENESIS.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED THROUGHOUT DEVELOPMENT.
 CC WHICH SHARES NO HOMOLOGIES WITH OTHER PTASE PROTEINS. THIS PTASE
 CC INSERT IS REMINISCENT OF THE KINASE INSERT WITHIN THE KINASE
 CC CATALYTIC DOMAINS OF SEVERAL RECEPTOR TYROSINE KINASES.
 CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PTASEIN-
 CC TYROSINE PHOSPHATASE FAMILY. SUBCLAS THAT CONTAINS SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: M94730; AAA28433.1; -
 CC EMBL: U19909; AAB02543.1; -
 CC EMBL: U19909; AAB02544.1; -
 CC EMBL: AE003423; AAF45724.1; -
 CC EMBL: AE003423; AAF45724.1; -
 CC EMBL: AE003423; AAF45724.1; -
 CC EMBL: AE003423; AAF45725.1; -
 CC EMBL: AL132797; CAB65870.1; -
 CC EMBL: AL132797; CAB65871.1; -

DR PIR: A43254; A43254.
 DR HSPB: Q06124; 2SHB.
 DR FLYBASE: FBgn0000382; CSW.
 DR InterPro: IPR003595; PTPC_motif.
 DR InterPro: IPR000980; SH2.
 DR InterPro: IPR000387; TYR_phosphatase.
 DR InterPro: IPR000242; TYR_PP.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00102; Y_phosphatase; 1.
 DR PRINTS: PR00700; PRTYPPHPTASE.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00194; PTPC; 1.
 DR SMART: SM00404; PTPC_motif; 1.
 DR SMART: SM00252; SH2; 2.
 DR PROSITE: PS00383; TYR_PHOSPHATASE_1; 1.
 DR PROSITE: PS00506; TYR_PHOSPHATASE_2; 1.
 DR PROSITE: PS00505; TYR_PHOSPHATASE_PTP; 1.
 DR PROSITE: PS00001; SH2; 2.
 DR Hydrolase; SH2 domain; Repeat; Developmental protein;
 KW Alternative splicing.
 FT DOMAIN 6 101 SH2 1.
 FT DOMAIN 111 205 -SH2 2.
 FT DOMAIN 227 645 PROTEIN-TYROSINE PHOSPHATASE.
 FT DOMAIN 289 444 PTPASE INSERT (CYS/SER-RICH).
 FT ACT_SITE 583 583 BY SIMILARITY.
 FT VARSPLIC 1 110 MSSRRFPTTISGIAEAKILODGFSGFLARLSSNGGAF
 TLYSRGNEVTHIKIIONNGDFDLGGEKFAFLPELYOYM
 ENGELKERNGOAIETKOPICAEPTTER -> MLEKLEK
 LSSILGANYNHLOEKQYNNNNNNNNNNNNNNNNNNNNQ
 RNFEYERAIQAHYSGKRRSERERSGFKASGKRAKAVTP
 PTEPEAEQACKNCMTDELAOLIKGVAKGADKARNDR
 LQRRRRPLSAQPSAASASTSTESLHRLTPQASAYPATPT
 SWTATPPQEPFAFGASCSNSTSLATMRVQLGHYT (IN
 ISOFORM 4A).
 FT FT CKMOCPAPLPRRGITILLISPVIFQONSKTFPRT -> A
 KFKNIPKDMIGLRPPSHAPALPPPTPPKRT (IN
 ISOFORM Y1229).
 FT FT G -> S (IN REF. 2).
 FT FT P -> S (IN REF. 2).
 FT FT 214FQF2576202CC CRC64;
 SQ SEQUENCE 845 AA; 92975 MM; 214FQF2576202CC CRC64;
 Query Match 6.8%; Score 136.5; DB 1; Length 845;
 Best Local Similarity 23.4%; Pred. No. 0.047;
 Matches 54; Conservative 39; Mismatches 67; Indels 71; Gaps 9;
 QY 170 SORHFEPEVOGMPQSIUSDLSLEVLAEKVPVHNO-----RKP 206
 DB 3 SRWHEPTISG-----LEAEKLQDQGDGSLARLSSNGCAFTLSVRGN 49
 QY 207 ESTHLENONTOEIPLAISSSFSTSNHVSQNRDHRGMOQPCSPQ--RCQPPASCSPHEN 264
 DB 50 EYTH-IKIONNGDFDLGGEKFAFLPELYOYMGELKENGALIEKQPLICA----- 104
 QY 265 ILPYKTSWRPPFKRSKDKDVQHNEMVIGEYSROAVEAFMKENKDGSLVYDCSTKSK 324
 DB 105 -----EPTTER-----WFGNLSGKEAEKLLIERGNGSFLVRE--SQSK 142
 QY 325 EEPYLAVEYENKYNVKIRFLERNQOFAITGIGDEKDEVEDIEIHYK 375
 DB 143 PDDEVLSVNTDKVTHVMIRW--QDKKIDVG-----GGESFTLSLIDHYK 187
 RESULT 11
 3BP2_HUMAN
 ID 3BP2_HUMAN STANDARD: PRT; 561 AA.
 AC P78314; O15373; O00500; P78315;
 DT 15-JUL-1998 (Rel. 36; Created)
 DT 15-JUN-1998 (Rel. 36; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE SH3 domain-binding protein 2 (3BP-2).
 GN SH3BP2 OR 3BP2 OR RES4-23.
 OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId:9606;
 RN [1]
 RP MEDLINE-97446134; PubMed-9299232;
 RX Bell S.M., Shaw M., Jau Y.-S., Myers R.M., Knowles M.A.;
 RA "Identification and characterization of the human homologue of
 RT SH3BP2, an SH3 binding domain protein within a common region of
 RT deletion at 4p16.3 involved in bladder cancer.";
 RL Genomics 44:163-170(1997).
 RN [2]
 RP SEQUENCE FROM N.A. (LONG AND SHORT ISOFORMS), AND TISSUE SPECIFICITY.
 RC TISSUE=Brain;
 RX MEDLINE-98403881; PubMed-9734812;
 RA Hadano S., Ishida Y., Ikeda J.-E.;
 RT "The primary structure and genomic organization of five novel
 RT transcripts located close to the Huntington's disease gene on human
 RT chromosome 4p16.3.";
 RL DNA Res. 5:177-186(1998).
 RN [3]
 RP SEQUENCE FROM N.A. (LONG ISOFORM).
 RC TISSUE=Tonsil;
 RA Gokemeijer J., Deligianidis K.E., Ligris K., Ernst T.J.;
 RT "3BP2 binds to phosphatidylinositols; linking the hemopoietic tyrosine
 RT kinase c-FES to the cytoplasmic membrane in a phosphorylation
 RT dependent mechanism.";
 RL Blood 88:473A-473A(1996).
 RN [4]
 RP VARIANTS CRBM GLN-415; PRO-415; ARG-418; HIS-418; LEU-418; ARG-420 AND
 RP GLU-420.
 RX MEDLINE-21275962; PubMed-11381256;
 RA Ueki Y., Tizabi N., Santana C., Fukui N., Maulik C., Garfinkle J.,
 RA Niomiya C., Domarai C., Peters H., Hahel N., Rhee-Morris L.,
 RA Doss J.B., Kreiborg S., Olsen B.R., Reichenberger E.;
 RT "Mutations in the gene encoding c-Abi-binding protein SH3BP2 cause
 RT cherubism.";
 RL Nat. Genet. 28:125-126(2001).
 CC -1- FUNCTION: BINDS DIFFERENTIALLY TO THE SH3 DOMAINS OF CERTAIN
 CC PROTEINS OF SIGNAL TRANSDUCTION PATHWAYS. BINDS TO
 CC PHOSPHATIDYLINOSITOLS; LINKING THE HEMOPOIETIC TYROSINE KINASE
 CC FES TO THE CYTOPLASMIC MEMBRANE IN A PHOSPHORYLATION DEPENDENT
 CC MECHANISM.
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
 CC short form: are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Expressed in a variety of tissues including
 CC lung, liver, skeletal muscle, kidney and pancreas.
 CC -1- DISEASE: Defects in SH3BP2 are the cause of cherubism (CRBM), an
 CC autosomal dominant inherited syndrome. It is characterized by
 CC excessive bone degradation of the upper and lower jaws, which
 CC often begins around three years of age. It is followed by
 CC development of fibrous tissue masses, which causes a
 CC characteristic facial swelling.
 CC -1- SIMILARITY: CONTAINS 1 PH DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: U56386; AAB72034.1; -
 CC EMBL: AB000462; BAA19119.1; -
 CC EMBL: AB000463; BAA19120.1; -
 CC EMBL: AF000936; AAB59973.1; -
 CC Genew; HGNC:10825; SH3BP2.
 CC MIM: 602104; -
 CC MIM: 118400; -
 DR DR
 DR InterPro: IPR001849; PH.
 DR InterPro: IPR000980; SH2.

FT DOMAIN 618 712 SH2.2.
SQ SEQUENCE 724 AA; 81059 MW; 9D2BABB6DE087098 CRC64;

Query Match
Best Local Similarity 21.2%; Pred. No. 0.062; DB 1; Length 724;
Matches 80; Conservative 47; Mismatches 122; Indels 129; Gaps 16;

QY 101 SIG-OPTMTQTRLEKVDKPSRDVRSQNIKGDASYRKKIKPLPPRLITLPPKTKQPLP 159
D 53 SVGMWGLNERTR-QRGDFPGT---YVEFLGPVALAR---PGRPRGRPLPARRDCP 104
QY 160 PEPESRPLPSORHTPEPQGMPSQISLSDSEVLEAEVPHNOKPESTHLEENOTOE 219
D 105 PEPGLTLPDLPEQFPSPDV-APPILVKT-----VEAIEPTGIDSYRPEPPAVTWSLSD 158
QY 220 IP-----LAISSTFTSNHSYONRDHRCGMQPCSPQ----- 251
D 159 VEQMDAALSDGVKGLLAPLPTPEAAAEHRLALRAAGVGPALPPLPLHALLT 218
QY 252 -----RCQPPASCP----- 261
D 219 LRFLLQHLGRVAGRAPAPGAVRATGATGCPILLRAPPPPPPGGAPDGTETPDPAL 278
QY 262 -----HENILPYVTSMRPPFPKRS-----DRKDVENWYIGETYSQ 299
D 279 LVEKLIQEHLEQEAAPAL-----PKRPKTPAPTGLANGSSPSIQDAEYMWGDISRE 334
QY 300 AVEAFMKENKDGSLFVRCSTKSKPEEYVLAFFENKYNYK-TRFLEKNOQFALGTL 358
D 335 EVNEK-LRTPDGTFLVRASSKIOGE-YTLVL---RKGGNKLKLVHRDGHYGESEPL 389
QY 359 RQDEKFDVEDIIEHYKN 376
D 390 ----TFCSYVDLITFHYRH 403

RESULT 13

P55G_HUMAN STANDARD; PRT; 461 AA.

AC Q92569; 060482; (Rel. 40; Created)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DT 16-OCT-2001 (Rel. 40; Last sequence update)
DE Phosphatidylinositol 3-kinase regulatory gamma subunit (PI3-kinase p85-gamma subunit) (Ptdins-3-kinase p85-gamma) (p55PIK).
GN PIK3R3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RA SEQUENCE FROM N.A. (ISOFORM 1).
RA Suzuki T.;
RT "Molecular cloning of human p55plk."
RL Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RC SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
RC TISSUE=Fetal brain;
RX MEDLINE=98192526; PubMed=9524259;
RA Dey B.R., Furlanetto R.W., Nisley S.P.;
RT "Cloning of human p55 gamma, a regulatory subunit of
RT phosphatidylinositol 3-kinase, by a yeast two-hybrid library screen
RT with the insulin-like growth factor-1 receptor."
RL Gene 209:175-183(1998).
CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
CC KINASES THROUGH ITS SH2 DOMAIN AND REGULATES THEIR KINASE
CC ACTIVITY. DURING INSULIN STIMULATION, IT ALSO BINDS TO IRS-1.
CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
CC SUBUNIT.
CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; 1 (SHOWN HERE), 2 AND 3; ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHEST LEVELS IN BRAIN AND TESTIS. LOWER
CC LEVELS IN ADIPOSE TISSUE, KIDNEY, HEART, LONG AND SKELETAL MUSCLE.

CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
CC -1- SIMILARITY: HIGH, WITH OTHER P85 SUBUNITS, AND WITH P85-ALPHA AND
CC P85-BETA SUBUNITS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: D88532; BAA13636.1; -
DR EMBL: AF028785; AAC39696.1; -
DR HSSP: P23727; 2PMB.
DR Genew: HGNC:8981; PIK3R3.
DR MIM: 606076; -
DR InterPro: IPR001720; PI3kinase_P85.
DR InterPro: IPR000980; SH2.
DR Pfam: PF00017; SH2; 2.
DR PRINTS: PR00678; PI3KINSEP85.
DR PRODOM: PD000093; SH2; 2.
DR SMART: SM00252; SH2; 2.
DR PROSITE: PS50001; SH2; 2.
KW SH2 domain; Repeat; Phosphorylation; Alternative splicing.
FT DOMAIN 34 44 PRO-RICH.
FT 65 160 SH2 1.
FT 358 452 SH2 2.
FT MOD_RES 341 341 PHOSPHORYLATION (BY SIMILARITY).
FT VARSPPLIC 36 71 MISSING (IN ISOFORM 3).
FT VARSPLIC 256 314 MISSING (IN ISOFORM 2 AND ISOFORM 3).
FT CONFLICT 21 21 P -> L (IN REF. 2).
SQ SEQUENCE 461 AA; 54462 MW; C03ECGF22BEB50B5 CRC64;

Query Match
Best Local Similarity 32.9%; Score 128; DB 1; Length 461;
Matches 49; Conservative 19; Mismatches 45; Indels 36; Gaps 10;

QY 240 DHRGMQPCSPQ-----RCQPPASCPHENILPYVTSMRPPFPKRS-----DRKD-----V 286
D 14 DRRVWMPYSTLILFIEMDPPA-----LPK-----PKPMTSAVPMCKDKSSVSL 60
QY 287 QHNEWYIGESYQAVEAEAPKENDGSLFVRCSTKSKPEEYVLAFFENKYNYK-TRF 345
D 61 QDAEYMWGDISREYNDK-LRMPDGTFLVRASIKMGD-YTLVL---RKGGNKLKIKI 115
QY 346 LERNOQFALGTGLRDEKFDVEDIIEHY 374
D 116 YHRDGKYGFSDDL-----TFNSVELLNHY 140

RESULT 14

CRK_DROME STANDARD; PRT; 271 AA.

AC Q9XYM0;
DT 16-OCT-2001 (Rel. 40; Created)
DT 15-JUN-2002 (Rel. 41; Last sequence update)
DE Adapter molecule Crk.
GN CRK OR CG1587.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RA SEQUENCE FROM N.A., AND CHARACTERIZATION.
RA TISSUE=Embryo.
RX MEDLINE=99173888; PubMed=10072777;
RA Galletta B.J., Niu X.-P., Erickson M.R., Abmayr S.M.;
RT "Identification of a Drosophila homologue to vertebrate Crk by
RT interaction with MBC."

RL Gene 228:243-252(1999).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN-Berkeley.
 RA MEDLINE-20196006; Pubmed-10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers J., Blake R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA April J.F., Agbayani A., An H.-J., Andrews-Planck C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostlin D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,
 RA Jaitai M., Kalush E., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclik J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheefer F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Styrbas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weissstock G.M., Weissensbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 CC -1- FUNCTION: ADAPTER PROTEIN WHICH INTERACTS WITH C-TERMINAL PORTION
 CC OF MBC, HOMOLOG OF HUMAN DOCK180. MAY PLAY A ROLE IN CELLULAR
 CC PROCESSES THROUGHOUT DEVELOPMENT.
 CC -1- TISSUE SPECIFICITY: EMBRYONIC ZYGOTIC EXPRESSION IS SEEN IN
 CC INVAGINATING PRESUMPTIVE MESODERM AND ECTODERMALLY DERIVED TISSUES
 CC DURING GASTRULATION. AT STAGE 8, EXPRESSION IS ALSO SEEN IN
 CC ANTERIOR AND POSTERIOR MIDGUT AND CEPHALIC FURROW. BY STAGE 9,
 CC EXPRESSION IS HIGHEST IN VISCERAL MESODERM OF ANTERIOR AND
 CC POSTERIOR MIDGUT, VENTRAL NERVE CORD AND SOMATIC MESODERM.
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY
 CC THROUGHOUT EMBRYOGENESIS, DECLINES DURING LARVAL STAGES AND
 CC REAPPEARS DURING PUPATION.
 CC -1- SIMILARITY: CONTAINS 1 SH2 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 2 SH3 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: AE003844; AA59362.1; -
 CC EMBL: AF112976; AAD28428.1; -
 CC HSSP: Q64010; 1CKA.
 CC FLYBase: FBgn0024811; Crk.
 CC InterPro: IPR000980; SH2.
 CC InterPro: IPR001452; SH3.
 CC Pfam: PF00017; SH2; 1.

DR Pfam: PF00018; SH3; 1.
 DR PRINTS: PR00401; SH2DOMAIN.
 DR PRINTS: PR00452; SH3DOMAIN.
 DR PRODOM: PD000066; SH3; 1.
 DR PRODOM: PD000093; SH2; 1.
 DR SMART: SM00252; SH2; 1.
 DR SMART: SM00326; SH3; 2.
 DR PROSITE: PS50001; SH2; 1.
 DR PROSITE: PS50002; SH3; 1.
 KM SH2 domain; SH3 domain; Repeat; Developmental protein.
 FT DOMAIN 12 114 SH2.
 FT DOMAIN 117 165 SH3 1.
 FT DOMAIN 220 259 SH3 2.
 SQ SEQUENCE 271 AA; 31205 MW; D1BAFEA3150932DC CRC64;
 Query Match 6.38; Score 127; DB 1; Length 271;
 Best Local Similarity 35.18; Pred. No. 0.054;
 Matches 33; Conservative 14; Mismatches 37; Indels 10; Gaps 3;
 Oy 281 SDRDVQHNENYIGEYSRQAVEAFMKENKDGSLVDCSTKSEPPYVLAVENRYN 340
 Db 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
 7 SDR-----NSWYFPMNRQADTEVLMNERGVFLVDSNSIAD--YVLCVREDTVSN 59
 Oy 341 VKIFLEPNOGFALGTGLRDEKDSVEDITEHY 374
 Db 60 YIINKVQODQDIYRI---GDQSFNDLPKLLTFY 90
 RESULT 15
 P85B_MOUSE STANDARD; PRT; 722 AA.
 AC 008908;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 15-DEC-1998 (Rel. 37, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phosphatidylinositol 3-kinase regulatory beta subunit (PI3-kinase
 DE p85-beta subunit) (Ptdins-3-kinase p85-beta).
 GN PI3K2.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NIH Swiss;
 RX MEDLINE-98241181; Pubmed-9582025;
 RA Janssen J.M.G., Schleithoff L., Bartlam C.R., Schulz A.S.;
 RT "An oncogenic fusion product of the phosphatidylinositol 3-kinase
 RT p85beta subunit and HWMORF8, a putative deubiquitinating enzyme";
 RL Oncogene 16:1767-1772(1998).
 CC -1- FUNCTION: BINDS TO ACTIVATED (PHOSPHORYLATED) PROTEIN-TYROSINE
 CC KINASES, THROUGH ITS SH2 DOMAIN, AND ACTS AS AN ADAPTER, MEDIATING
 CC THE ASSOCIATION OF THE P110 CATALYTIC UNIT TO THE PLASMA MEMBRANE.
 CC -1- SUBUNIT: HETERODIMER OF A P110 (CATALYTIC) AND A P85 (REGULATORY)
 CC SUBUNIT.
 CC -1- SIMILARITY: BELONGS TO THE PI3K P85 SUBUNIT FAMILY.
 CC -1- SIMILARITY: CONTAINS 2 SH2 DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 SH3 DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: Y13569; CA73903.1; -
 CC HSSP: P23727; 2PNB.
 CC MGD: MGI:1098772; PI3K2.
 CC InterPro: IPR001720; PI3K_kinase_P85.
 CC InterPro: IPR000198; RHOgap.

DR InterPro: IPR000980; SH2.
 DR InterPro: IPR001452; SH3.
 DR Pfam: PF00017; SH2; 2.
 DR Pfam: PF00018; SH3; 1.
 DR Pfam: PF00620; RhogAP; 1.
 DR PRINTS: PRO0678; PIKINASEP85.
 DR PRINTS: PRO0401; SH2DOMAIN.
 DR ProDom: PD000066; SH3; 1.
 DR ProDom: PD000093; SH2; 2.
 DR SMART: SM00324; RhogAP; 1.
 DR SMART: SM00252; SH2; 2.
 DR SMART: SM00326; SH3; 1.
 DR PROSITE: PS50001; SH2; 2.
 DR PROSITE: PS50002; SH3; 1.
 DR SH3 domain; SH2 domain; Repeat.
 KW SH3 domain; SH2 domain; Repeat.
 FT DOMAIN 4 80 SH3.
 FT DOMAIN 125 257 RHO-GAP.
 FT DOMAIN 324 419 SH2 1.
 FT DOMAIN 616 710 SH2 2.
 SQ SEQUENCE 722 AA; 81251 MW; AB4D49AF30CEC567 CRC64;

Query Match

Best Local Similarity 6.3%; Score 127; DB 1; Length 722;

Matches 100; Conservative 50; Mismatches 146; Indels 129; Gaps 28;

OY 7 PNRSM-PRINSAT-----GQYQRMNKPL---LDWERNFAAV 39
 DB 51 PHNVGMMPGEMERTROGDFCTYVEFLGPVALARGPRGRPLPARPLDGSSSESHI 110
 OY 40 L-DGAKGSHDDYDDELMETWOSIKILPARPIKESYADTHYKVMATPLPLDTRT 98
 DB 111 LFDLAEOGSPPPAPPIL-----VKLYEA--IEQAEIDSECYSK---PELPA-TRT 155
 OY 99 SISIGOPTWNTQRLERVVKPISRDVRSQNIKGDAVSRKNKILPPEPRPLIT---LPKKY 155
 DB 156 DMSLSD-----LEQMDRTALYDA---VKG-----FLALPAAVTPEAAAEAY 195
 OY 156 QPL-----PEPESSRPPLSQRHTFPEVOGMPQSILRLSEVLEAEKYPHNQKPE-STH 210
 DB 196 RALREVAGVGLLEPP-----TLPLHQALTLRFLQLHGRV--ARRAP-----SPDTAVH 244
 OY 211 LL-----ENQTOELP-LAISSSFTTSNHSVQNDHRGGMQPCSPQRC 253
 DB 245 ALASAFGLLRIPPSGGEGDSEVPDPFVILLERLYOEH--VEEQDAAPALPPKPSKA 303
 OY 254 QP-PASCSPHENILPYKYSWRPPPKRSRDKVOHNEMYIGEYSROAVEAFMKENKDG 312
 DB 304 KPAFTALA-----NGGSP-----SLQDAEMTWGDISREYNER-LRDTPDG 344
 OY 313 SELVDCSTKSEKPYVLAIFYENKVVYVK-IRFLERNOQAFALGTGRGDEKFDSEVEDII 371
 DB 345 TFLVWDASKIGE-YTLTL--RKGNNKLKVFHRDGHYGFSEPL---TFCSVVELI 396
 OY 372 EHYKN 376
 DB 397 SHYRH 401

Search completed: April 21, 2003, 12:36:55
 Job time : 12.3453 secs

